TILBURG UNIVERSITY

Tilburg University

Essays on public policy and household decision making

Kabátek, Jan

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Link to publication in Tilburg University Research Portal

Citation for published version (APA):
Kabátek, J. (2015). *Essays on public policy and household decision making*. CentER, Center for Economic Research.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

JAN KABÁTEK

ESSAYS ON PUBLIC POLICY AND HOUSEHOLD DECISION MAKING

ESSAYS ON PUBLIC POLICY AND HOUSEHOLD DECISION MAKING

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in aula van de Universiteit op dinsdag 30 juni 2015 om 14.15 uur door

Jan Kabátek

geboren op 22 augustus 1987 te Havlíčkův Brod, Tsjechië.

Promotor
Prof. dr. Arthur van Soest

Promotiecommissie
Prof. dr. Jaap Abbring
Dr. Hans Bloemen
Dr. Mauro Mastrogiacomo
Prof. dr. Andreas Peichl

Dedicated to Václav Polák.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

Chapter 2
Kabátek, J., A. van Soest, and E. Stancanelli (2014): "Income taxation, labour supply and housework: A discrete choice model for French couples" *Labour Economics*.

Chapter 3
Apps, P., J. Kabátek, R. Rees, and A. Van Soest (2012): "Labor supply heterogeneity and demand for child care of mothers with young children," *IZA Discussion Paper* 7007.

Chapter 4
de Boer, H.-W., E. Jongen, and J. Kabátek (2014): "The effectiveness of fiscal stimuli for working parents," *CPB Working Paper* 116.

ACKNOWLEDGMENTS

A learning experience is one of those things that says, 'You know that thing you just did? Don't do that.'

Douglas Adams

Before I indulge in the very purpose of these lines, I have to admit that I approached writing the Acknowledgments with deliberate caution and a vague sense of dread. Being fully aware that it is likely to be the only part of my thesis which will have readership wider than the members of my dissertation committee, I felt pressured to craft each of its sentences into a solid nugget of wisdom and reflection. This, as you are soon to realize, proved infeasible and for that I apologize. Conveniently, I blame the time constraints.

Six years ago, I fell through the rabbit hole and started exploring the strange world of academia. During my journey I met many incredible characters (often as bizarre as the ones concocted by Lewis Carroll), traveled far and wide, and learned quite a bit about my discipline and myself as well (usually following the template described in the opening quote). Throughout this time, I was fortunate enough to be accompanied by a host of incredible mentors and friends to whom I would like to express my sincere gratitude.

I cannot but start with my supervisor. Arthur, we have been through many things together - writing papers, setting up a graduate econometrics course, traveling, eating *a lot* of Thai food, and also that one time when we waded through a river while looking for koalas. But most importantly, you gave me the opportunity to start working on a sound empirical research project while still being in my first year of coursework. You may not be aware of this, but being your research assistant became a turning point of my life. It lifted me from below the poverty line, it kept me passionate about economics despite the drudgery of graduate coursework, and most importantly, it showed me the profound appeal of policy-relevant quantitative analysis that led me to pursue the doctorate as such. For that I am deeply thankful. During all these years you have been an incredible supervisor, providing me with a plethora of time and resources, and teaching me great deal about research, econom(etr)ics, and academic work in general. You are and always will be an example to follow.

I want to thank the members of my doctoral committee, Andreas, Hans, Jaap and Mauro. I greatly appreciate their expertise and the time and effort they put into their committee service. They provided me with many great insights and comments that turned the papers underlying this thesis into a robust and inclusive manuscript which I am very happy about. The papers themselves would however not be

here without my coauthors - I want to thank Arthur, Egbert, Elena, Henk-Wim, Patricia and Ray, who showed me many facets of economics, taught me a lot about the collaborative process, and corrected all-too-many definite articles which I misplaced due to my Slavic predicament.

I want to thank Martin Salm for being willing to listen to me whenever I had something on my mind. The sincerity of his advice and his keen interest in the welfare of our graduate school are nothing but laudable. The job market committee (Otilia in particular) deserves a great credit for their active engagement in the final stretch of my studies. Big thanks goes to Johannes for his hospitality and his healthy opinions on economic research and education (and climbing technique). I also want to express my gratitude to Meltem who told me to get out of my comfort zone and interact with seminar speakers on a regular basis. Indirectly, she is responsible for my visit to the University of Pennsylvania where I spent one of the best semesters of my life. Regarding the Penn faculty, I greatly benefited from discussions with Andrew Shephard, Frank DiTraglia, Ken Wolpin, and my host, Petra Todd, whose council was particularly enlightening.

I had two amazing officemates, both very quiet, yet of intellect far beyond my comprehension. The first was Juan Juan with her *extreme* research and the most adorable Chinese-Italian accent, and the second was Yifan with his voracity for knowledge and sophisticated post-seminar discussions. I want to thank both of them for enduring my often not-so-quiet presence.

Needless to say, there were many other friends who made my days brighter and less stressful. I want to thank Hanka and David, my Czech enclave in Tilburg and (highly amusing) second family, Luc, the kindred (and kind) spirit hiding under a shell of stubborn pessimism, Michele, my soft-spoken partner in mischief, and Jan, my favourite teammate, trusty Penn connection and incessant resource of sarcasm. I will cherish the moments spent with Sara and Stefan, who helped fueling my caffeine addiction, Bas, who made sure I didn't work too long without a break, Renata and Sebastian with whom I engaged in long discussions (albeit on very different topics), the 'old' lunch group (Dominik, Gaia, Jarda, Marco, Rasa, Yan & Ying), the 'new' lunch group (Alaa, Gyula, Krzysztof, Maria, Marieke, Marleen & Nick), Ivo and Edith, the highly-praised cakonometrics group, the sauna club, and the climbing team. Finally, I am grateful to the handful of people I could lean on when I needed it the most, Elizabeth, Larissa, and my true best friend, Yuri.

A special place in my heart is reserved for my family, who stood by me as a beacon of certainty, supporting me at every step I took. Thank you for your unconditional love, there is nothing I hold dearer.

Jan Kabátek
Tilburg, May 17, 2015

CONTENTS

1

INTRODUCTION

Public policy systems around the world are changing at an ever-increasing pace. Taxes and benefits are being introduced, adjusted, and discarded in order to stimulate the economy, correct for market imperfections, or promote redistributive goals of the government. Many of these policies are directly targeted at couples and larger families, constituting what is often called *family tax policy*. This part of the tax system focuses on the issues which are specific to multi-person households. The family tax policy debates are traditionally under careful public scrutiny, because their topics touch people's daily lives. The issues discussed are diverse, including but not restricted to: Should we tax the incomes of both spouses separately, or should we add them up and apply a single tax rate? Should we promote institutionalized child care, or should mothers be the ones looking after their children? Should we promote fertility of the population while sustaining active female workforce, or is it preferable to have mothers staying outside of the labor markets?

Albeit being highly selective, this list illustrates that one of the main discussion points of current debates is female labor supply. Many countries are actively trying to promote female labor supply, introducing new policies and reforming the old ones in order to increase female labor participation and narrow down the gender wage gap. But when looking at such interventions, we should ask: How effective are these policies in reaching the desired goals? Could we do better if we spent more money on the reforms? Or should we rather consider alternative policy proposals? These are the questions which can be answered through the prism of structural modeling, which is the unifying theme of the chapters of this thesis.

The analysis of household decision making is a complex domain of economic modeling which deviates in many ways from the standard textbook approaches of labor economics and public economics. The economic model which is traditionally used in these fields is the one of a single decision-maker, who divides her time between leisure and labor. Leisure is desirable since it increases individual utility, but it is not a productive use of time (at least in the context of this model). Labor requires effort, which lowers individual utility, but this loss is compensated by income. Income increases utility since it can be exchanged for goods, and it generally renders some form of labor desirable. For that reason, the decision maker divides her time between labor and leisure time in a way that would maximize her well-being. Needless to say, this workhorse model has proven extremely useful

for the development of labor economic theory, yielding many insights about individual economic behavior. However, due to its inherent simplicity, it fails to account for important aspects of the household decision making which are crucial for our analysis.

In partnered households, we have two persons whose choices are mutually dependent, and whose utilities are fundamentally intertwined. Keeping the primitives of the problem similar to that of the workhorse model, this means that the partnered households are facing two interdependent labor/leisure trade offs. Both spouses have to decide how much they want to work, which complicates the analysis due to all the possible labor & leisure combinations that can be chosen. The joint nature of the problem is however beneficial for the household, because spouses can now substitute each others' labor supply (assuming they pool their earned incomes). That way, the couple can exploit comparative earnings advantage of the more productive spouse, increasing the joint well-being of the household.

The analysis becomes more involved if we allow for alternative time-use choices such as non-market work (which is regarded as leisure in the workhorse model). The rationale for separating non-market work from leisure follows from the premise that the spouse who substitutes away from working in the market is not going to spend her time entirely on leisure activities. Rather than that, she will specialize in non-market work which, unlike leisure, is a productive use of time generating *household good* (a blanket term for all the goods and services provided for the family outside of the formal labor markets). This distinction is of considerable importance to the analysis of female labor supply. Once a spouse decides to start working in the market, she has to forgo not only her leisure time, but also her time spent on non-market work. With less time in the household, her own provision of household good is going to fall. But since the household good tends to be highly desirable, the family will try to obtain it from alternative sources - either by delegating other household members or by buying it on the market. And as long as there are exercisable alternatives for household good provision, the woman will enjoy sufficient flexibility in her work choices. However, in the households which do not enjoy enough leisure time and who cannot afford to pay for market provision, the work choices of women are going to be rather limited. This can effectively prevent women from engaging in market work (which we would not be able to capture without making the distinction between leisure and non-market work). Accordingly, the inclusion of non-market work in the model can have considerable impacts on the resulting analysis, being particularly important for evaluating policies targeted at low-skilled female labor supply. This dependence is documented in Chapter 2, which allows for a model where spouses choose both market work and non-market work allocations.

The non-market work becomes ever so important with the arrival of children, as spouses have to dedicate substantial part of their time to child care. The amount of time required by children is usually decreasing as the children grow older, but especially in the pre-school period it is likely to impose considerable constraints on parental time. The need for pre-school child care provision often keeps women from engaging in market work, unless they decide to substitute their own care by the service of others. This, in turn, depends on several factors, including availability of informal care providers (such as grandparents or other relatives), availability of formal care providers (child care centers and kindergartens), maternal preferences for own child care provision, and other household characteristics such as the financial situation of the family. All these factors contribute to the mother's decision whether to engage in the market work or not, and it can be expected that the observed households will exhibit substantial heterogeneity across all of the mentioned domains. Therefore, in order to capture women's attitudes towards market work, we have to pay careful attention to modeling their child care preferences, access to different modes of care, and its dependence on various household characteristics. Accordingly, in Chapters 3 and 4 we employ models of household decision making which allow for joint choice of labor supply and child care.

Having acknowledged that the decisions of spouses are critically dependent on having children, the next step in the analysis of household labor supply is to focus on the actual arrival of children, that is, modeling fertility. This extension is pursued in Chapter 5. Accounting for fertility becomes particularly important when evaluating policies which are likely to impact not only the female labor supply, but also childbearing (*e.g.,* maternity leave or childcare subsidies). Following the observation that some couples have children at a very young age, whereas others postpone childbearing to much later stages of life, we can attempt to capture their chosen life paths by modeling fertility as a choice. The choice to bear children can be related to preferences, earnings potential, current (or expected) financial situation, and other household characteristics. That way, we can identify the fertility effects for various subsamples of the population and evaluate the proposed policies in a much more comprehensive framework.

All these extensions are trying to account for endogeneity present in the original workhorse model. We show that the "leisure" category is too broad to be treated as a homogeneous good which drives the decisions of the whole population in a uniform way. Firstly, the introduction of non-market work and the implied necessity of household good provision differentiates between the work options of well-endowed households and those living on a budget, pointing out that we cannot expect the same degree of work flexibility within both groups. Secondly, accounting for heterogeneity in the preferences for child care

will refine the role of non-market work for labor supply decisions of parents with young children. The last extension illustrates that there are further decisions taken by the households which are likely to be influenced by fiscal policy. Allowing families to have more (or fewer) children in response to changing incentives brings the model closer to reality. The corresponding reform predictions are therefore able to capture not only the effects for families whose composition remains stable, but also for the families who decide to have a child (or more children) as a result of the proposed policy change.

This dissertation contains four core chapters which follow through with the extensions discussed above. The ordering of the chapters reflects the chronological order in which the corresponding papers were appearing. I also like to think that it reflects the five-year-long organic development of our ideas and research agendas, starting from relatively simple models and moving to more involved, comprehensive analyses of household decision making. This shift is reflected both in terms of the complexity of our models and the econometric methods used to estimate them.

From the policy perspective, the thesis can be divided into two thematic parts. The first part consists of two chapters which are focused on the effects of joint versus individual income taxation of couples in France and Australia, respectively. The policy analyses in both papers are targeted at quantifying the labor supply responses and revenue effects induced by a shift from joint income taxation to individual income taxation. The last two chapters analyze the fiscal stimuli for working parents, focusing on child care subsidy and tax credit reforms in the Netherlands. The policy goal of both chapters is to determine which of the policies - the child care subsidy or the tax credit - was more effective in stimulating the female labor supply. This is assessed firstly in a static, and subsequently also in a dynamic modeling framework.

Part I: Moving from joint to individual income taxation

Income taxation constitutes one of the most robust pillars of public policy systems around the world. Workers pay income tax by remitting part of their earnings to the tax authority, and the size of this remittance depends on the applicable income tax rate. The income tax rates are generally increasing with income, which means that the poor are taxed less than the rich, both in absolute and in relative terms. This favorable treatment of low-income households constitutes what is known as progressive income taxation.

One of the key questions in family tax policy is how should income tax legislation treat couples. Generally, we distinguish two specific tax regimes: the spouses can be either regarded as one entity with a single tax rate applicable to their combined incomes, or they can be taxed

separately with individual tax rates based on their respective incomes. These two systems are known as joint income taxation, and individual income taxation. Nowadays, the majority of OECD countries has adopted individual income tax systems, but a handful of countries (*e.g.,* Australia, France or Germany) are keeping the joint systems in place.

The joint income taxation has been repeatedly criticized by Apps and Rees (1988, 1999, 2011) for its adverse effects on female labor supply. With the joint income tax in place, women in partnered households become less likely to join the labor force, since as the second earners they are facing high income tax rates irrespective of their work efforts. The disincentives for labor participation become particularly apparent if we compare market work (yielding highly-taxed earned income) with non-market work in the household (yielding household good which is not burdened by any tax). Many women will therefore substitute their market work by the non-market work, taking advantage of the favorable tax treatment of household good production.

In the two chapters which fall into this part of the thesis, we explore the dependence of household decision making on the income taxation regime in place. More specifically, we show what would happen if France and Australia, two of the countries that are currently using joint income tax system, switched to individual income taxation.

Chapter 2 analyzes the joint work and housework decisions of married and cohabiting couples in France. It takes advantage of the French Time Use Survey, a household-level dataset which contains detailed information on the ways how French couples allocate their time between work, housework and leisure. In order to capture work and housework preferences of the surveyed households, we develop a structural model of household decision making, following the seminal work of Van Soest (1995). The model used is a static discrete choice model which accounts for observed and unobserved heterogeneity in households' preferences. The unobserved heterogeneity is introduced by allowing individual preference parameters to be drawn from a specified distribution, rendering the utility function semi-parametric.

We estimate own wage elasticities of labor supply for both spouses, finding that women are much more elastic in their labor supply than their male partners. This result holds also if we look at cross-wage elasticities, letting the spouses react to changes in wages of their significant other. The core of the paper is dedicated to the analysis of shift from joint to individual income taxation. We show that the joint taxation indeed discourages female labor supply, with women's working hours being 3.7% higher under the individual taxation regime. The switch to individual income taxation is also associated with a fall of women's housework hours by 2.0% on average. Men's responses

are found to be less pronounced, which is in line with their lower labor supply elasticities. On average, men are predicted to reduce their market work hours by 0.8% and increase their housework hours by 1.3%. The increase in men's housework hours is however not big enough to cover the entire drop of women's hours. This means that there is less household good produced by the spouses themselves, suggesting that the production is likely to be outsourced to the market providers.

Chapter 3 builds on the groundwork set out in Chapter 2, shifting the focus to a more specific population of interest - mothers with pre-school-aged children in Australia. As outlined above, mothers with young children are likely to be more restricted in their work choices. The role of non-market work will be emphasized in their decision making, since children require a substantial amount of attention and care, leaving aside all the other forms of necessary housework. This is confirmed by cross-country evidence presented in Apps and Rees (2009), showing that female labor participation falls drastically in the years after first childbirth. The resulting gender gap in labor participation narrows as children grow older, but it never completely disappears.

The decline of women's work hours is however far from uniform. Mothers with young children are likely to exhibit very different attitudes towards market work depending on the availability of informal & formal child care providers, and other factors mentioned in the prior discussion. For the sake of sound policy analysis, this means that we have to pay particular attention to child care when modeling maternal labor supply. In this chapter, we benefit from a very rich household survey of Australian population (HILDA), which contains meticulous information on different modes of child care that may be used by the surveyed families. The information on formal and informal child care utilization is directly incorporated into our model. The model is again a structural discrete choice model, but this time it considers a different set of choice variables: mother's labor supply, mother's non-market work engagement, and formal child care use. Here, the major difference compared to Chapter 2 is that the father's behavior is treated as given. The only way how the father can influence outcomes of the model is through his income (which is supplied to the household). This assumption reflects stricter gender roles in Australian households, where men are more often the sole breadwinners whose choices are unlikely to respond to the decisions of their partners (see Doiron and Kalb 2005). It also allows us to limit the computational complexity of the baseline model, which allows us to pursue more flexible treatment of unobserved heterogeneity. Following Train (2009), we adopt a latent class model, which uses different distributional assumptions on the individual preference parameters and which is

generally considered to be more flexible than the random coefficient method used in Chapter 2.

We estimate own wage elasticities of labor supply for mothers with pre-school children, showing that they are slightly lower than the ones obtained from the sample of French women. Australian mothers are found to reduce their market work if they are subject to an increase of child care prices, substituting the expensive service by their own child care provision. We also show that unobserved heterogeneity in preferences is playing crucial role in household decision making and failure to account for it results in biased preference parameters and misleading policy implications. Similarly to the previous chapter, we investigate the effects of installing purely individual income taxation in Australian fiscal system, changing all the taxes and levies which are dependent on joint income of the household. Our findings confirm the conclusions of Chapter 2 - women are more likely to engage in market work in the system which is based on individual, rather than joint income taxation. Predicted increase in mothers' market work hours is 3.41% on average, which is accompanied by 2.74% increase of average formal child care hours.

Part II: Effectiveness of fiscal stimuli for working parents

It should be noted that the choice of population of interest in the previous chapter is by no means accidental. The focus on mothers with young children follows from the recent academic and policy debates which emphasize active engagement of mothers in the labor force. The labor participation of mothers with young children has been identified to be highly influential for women's later-life work choices (see Bernal 2008, Bernal and Keane 2011), implying that women are likely to follow through with the work habits formed in the child-rearing phase. This finding is important, since it makes way for targeted policies which have the potential to be highly effective in stimulating female labor supply whilst bearing relatively low budgetary costs.

In other words, we should incentivize the labor supply of mothers with young children, because they constitute a relatively small group within the population and they are likely to stick to their working arrangements for many years to come. Naturally, this incentivization can take on many forms. Governments can support working mothers financially, provide them with occupational training, or facilitate their access to the labor markets. In the following chapters we focus on the policies using financial incentives which are also known as fiscal stimuli.

The fiscal stimuli act by means of providing mothers with conditional financial transfers to make their market work engagement more attractive. Often, these transfers take on the form of child care subsidies, following the rationale that some mothers would opted for

market work had they have sufficient means for buying-in formal child care. Alternatively, governments can employ in-work tax credits which provide additional income to all parents who start working in the market.

Needless to say, there are pros and cons to both child care subsidies and tax credits. Child care subsidies are likely to stimulate both female labor supply and the child care service sector, making formal child care services easier to access. On the other hand, the subsidies are criticized for crowding out the informal child care sector: with the subsidies in place, a family with two employed spouses may opt for formal child care, substituting the informal service provided by grandparents or other relatives. However, since both spouses are already working, they are unlikely to increase their market work hours further. This substitution of modes of child care therefore creates an allocative inefficiency which is hard to circumvent without intrusive oversight and excessive targeting of child care subsidies.

In-work tax credits are generally considered effective in stimulating female labor supply, but their outcomes may also go awry. For example, the American version of the credit, the Earned Income Tax Credit (EITC), has been found to have strong negative effects on labor supply of mothers in partnered households (Baughman and Dickert-Conlin, 2003). This outcome has been attributed to the eligibility criteria for EITC, which require only one of the spouses to be actively working. As a result, women in the position of secondary earners face lower incentives to work, since the tax credit is awarded to the family irrespective of their own labor engagement.

All things considered, it is still hard to determine which of the policies would prove more efficient in stimulating the maternal labor supply. This tasks calls for a structural model which would help to uncover the relative effectiveness of child care subsidies and in-work tax credits (and their various types). We investigate this issue in two alternative modeling frameworks, starting with a static model and then moving into a dynamic setting. Each analysis constitutes a separate chapter of this part of the thesis.

Chapter 4 presents a static model of household decision making which borrows parts of the modeling setup from the Chapters 2 & 3. We analyze the population of married and cohabiting adults in the Netherlands who have children younger than 12 years of age. The dataset comes from national administrative statistics (Statistics Netherlands), and it provides us with detailed and high-quality data on work & child care choices for a sizable fraction of the Dutch population.

The fact that our analysis is set in the Dutch context has several modeling implications. On one hand, we can take advantage of large-scale

policy changes which happened in the country in the observed period: Throughout the years 2004–2009, the Netherlands implemented several reforms of child care subsidies and in-work tax credits for families with young children, rendering maternal market work much more attractive. These changes of credits and subsidies represented a sizable (and arguably exogenous) variation in the incentives underlying the household decision making. And since we know how the programs actually changed, we can incorporate them explicitly into the model. By doing so, we allow the structural parameters to be partially identified by the exogenous variation in the incentives, and thus we are likely to improve on the inference of our model.

The Dutch context however brings also some complications. In the Netherlands, the archetype of a sole male breadwinner is much less prominent than in the Australian households, and therefore it seems more appropriate to allow the work choices of both spouses to be jointly determined. Accordingly, the modeled decisions include: labor supply decision of men, labor supply decision of women, and the decision to use formal child care. In contrast to the previous chapters, the non-market work is not treated as a separate choice. It enters the model combined with pure leisure time in a composite "leisure" indicator. This treatment does not allow us to analyze the effects of the reforms on intra-household allocation of housework, but it is necessary since the administrative statistics lack the information on spouses' non-market work choices. The preference heterogeneity is treated in the same way as in Chapter 3, using the latent class model to explore the importance of unobserved factors in the utility function.

We estimate own wage elasticities of labor supply for both spouses, confirming the French result that women are more elastic in their work choices than men are. The elasticities for women are on par with those extracted from the Australian data. The elasticities for men are lower than those attributed to French men, however both these results should be taken with a grain of salt, since the French elasticities correspond to more heterogeneous population sample. The policy analysis focuses on assessing the relative efficiency of several parameterizations of child care subsidy and in-work tax credit reforms, including the ones that were actually implemented in the Netherlands. We find that in-work tax credits that are targeted at second earners are the most cost-effective instruments for raising the female labor supply. The difference between the two implemented policies is however found to be rather small. An interesting outcome of the analysis is that the effectiveness of both policies could be further improved if they were set to increase with second earner's income.

Chapter 5 is in many ways similar to Chapter 4. We ask the same policy questions, we explore the same data source, and both the analyzed households and the periods of observation are to a large

extent overlapping. There are, however, several important factors which distinguish Chapter 5 from the previous analyses.

First of all, the static model which characterized all the works presented so far has been replaced by a dynamic model. The dynamic structural model treats the decision maker as a forward-looking agent. This means that her choices are reflecting the best option for both her current and her future selves. This distinction becomes particularly important when there is a direct link between her current choices and the choices in the years to come. In the context of labor supply choice, this link is embodied by human capital accumulation. By working in the market, a woman will accumulate experience which will increase her wage and improve her standing on the labor markets in the years to come. Thus, her decision to work now is likely to influence her work decisions down the road, since the gained experience will change the *state* in which she finds herself when making the next year's work decision.

Accounting for human capital accumulation is very important for assessing the long-run impacts of the competing fiscal stimuli. In the short-run, the static model may prove to be informative enough, but after 5 or 10 years, the cumulative effects of work experience are likely to play in resulting outcomes. The human capital link may also represent one of the factors explaining why we see strong persistence of maternal labor supply choices. If it is the recent experience that matters the most for women's decision making, then the loss of experience in the periods of early child rearing can induce highly persistent gaps between male and female labor supply for many years after the childbirth.

The second factor which distinguishes Chapter 5 from the others is the treatment of fertility. Since the previous models were static, the number of children in the family was implicitly kept fixed at the observed levels. In this model, the fertility is considered to be a choice, entering the household's choice set together with women's labor supply and formal childcare choices. This allows us to account for the fact that introducing financial incentives for working parents may enhance fertility, because prospective childbearing becomes less costly. This channel is an important ingredient for the study of long-run implications of the reforms targeted at working parents. If a policy, such as child care subsidy, is likely to stimulate fertility, then the costs of the policy are likely to increase in the years after the implementation as there will be more families and more children eligible for the subsidized care.

Our empirical model builds on the current developments in the field of dynamic female labor supply, particularly on the works of Francesconi (2002) and Bernal (2008). The labor choices in the household's choice set are once again restricted to women, treating men's labor supply as given. In the Dutch context, this is more stringent

assumption than in Chapter 3, however such restrictions are necessary to limit the computational complexity of the model. Furthermore, we have showed in Chapter 4 that men's work choices are not too responsive to changes in the incentives faced by the households. The preference heterogeneity is restricted to the inclusion of several observed characteristics into the utility function, and making the unobserved components corresponding to each choice variable correlated among themselves.

The core of the paper is policy analysis which assesses relative effectiveness of the 2004–2009 reforms in the Netherlands. Their effects are analyzed both in the short run, *i.e.*, immediately after the implementation, and in the long run, which denotes 10 years from the implementation. In the short run, the results are very similar to those found in Chapter 4. The two implemented reforms are shown to be similarly cost-effective, with the difference between the two policies being very small. However, in the long run, the child care subsidies become much more costly due to the effects child care subsidization has on fertility. The maintenance costs of the childcare subsidies are projected to rise by more than 40% over the 10-year period, whereas the costs of in-work tax credits are predicted to fall due to the effects of human capital accumulation. This result is supportive of the claim that in the short run static analyses are likely to produce reasonable approximation of the reform effects. On the other hand, it also shows that these analyses fall short on capturing the dynamic changes which are initiated by the reforms but unlikely to manifest earlier than several years after the implementation.

A note to the interested practitioner

This dissertation is a collection of empirical analyses which all investigate similar phenomena, and as such it can be of interest to the practitioners who want to explore the domains of household decision making and public policy analysis. The following chapters, albeit thematically related, however differ in many of the imposed modeling assumptions which raises the question, what is the most appropriate set of assumptions to adopt when analyzing the household decision making?

Unfortunately, there is no easy answer to this question. The differences in assumptions which we adopt in each chapter are to a large extent idiosyncratic - they reflect the research questions we set out to answer, the cultural and institutional background of the populations of interest, the limitations of our data, and the computational complexity of the estimation methods we use. And even though tt is true that the ordering of chapters reflects the relative complexity of the employed models and the methods used to estimate them, it would be misleading to claim that the assumptions used in Chapter 4

are more preferable than the assumptions used in Chapter 3. Both reflect the specific context of the pursued analyses, and this context-dependence is likely to be present in any given empirical investigation of household decision making.

In some cultures, the assumption postulating that men do not respond to their partner's labor choices will be fully justifiable, whereas in other cultures it will not. Similarly, in some contexts we can abstract from modeling the fertility decisions, and in others we cannot. The practitioner has to decide which assumption is appropriate in the current context, based on information and resources at her disposal. This can be done by allowing for higher flexibility of the model and testing the assumption in question. However, doing so can often prove infeasible due to data limitations or other estimation issues. In such situations, we have to make a judgment call, assessing the validity of given assumption in light of our economic intuition or anecdotal evidence.

As you will see, both of these approaches are exercised many times in the chapters of this manuscript. I have made my best effort to cross-reference individual chapters, highlighting the differences in their respective assumptions. Often, the maintained assumptions are assessed using various robustness checks. Other times, we discuss potential extensions of our model, pointing out how would the results change if the maintained assumptions proved unjustified. I hope that this exposition will reflect both the merits and the limitations of standard assumptions in the structural models of household decision making, and that it will serve the interested practitioners as a useful guide for their own model building.

Part I

MOVING FROM JOINT TO INDIVIDUAL INCOME TAXATION

2

INCOME TAXATION, LABOUR SUPPLY AND HOUSEWORK: A DISCRETE CHOICE MODEL FOR FRENCH COUPLES

This chapter is the reproduction of a paper written with Elena Stancanelli and Arthur van Soest, published in the *Labour Economics*.

2.1 INTRODUCTION

Theoretical studies of income taxation conclude that income taxes may affect not only individual labour supply but also the amount of domestic work produced within the household. Income taxation is likely to affect labour supply and housework hours in opposite directions, because, for instance, downward changes in the individual rewards from work reduce the individual opportunity cost of housework and thus, housework becomes more attractive than market work. There is limited empirical evidence on this issue. This paper adds to the literature by estimating a discrete choice model of both partners' market and housework hours. Using these estimates, we simulate how a change from joint to separate taxation of married spouses' incomes affects spouses' hours of market and non-market work. This is especially interesting since France is one of the few OECD countries that still taxes the incomes of couples jointly.

Apps and Rees (1988, 1999, 2011) argue that although household production is not taxed (which is unavoidable since its output cannot be observed), the taxation of market work is likely to affect housework hours of spouses and, in particular, married women's labour supply is likely to increase when replacing joint taxation by separate income taxation.[1] Leuthold (1983) estimated the tax elasticities of housework of husband and wife in one and two-earner US households using a single equation framework, and found that (joint) income taxation increases housework done by women and reduces housework done by men. Gelber and Mitchell (2012), focusing on American single women, concluded that when the economic rewards for participating in the labour force increase, single women's market work increases and their housework decreases. Rogerson (2009) examined the effects of taxation on housework and labour supply in the US and Europe from a macroeconomic perspective, and found that when accounting for

1 See also Kleven et al. (2010) for a recent treatment of the optimal taxation of couples. Alesina, Ichino, and Karabarbounis (2011) analyze how "selective" taxation, i.e., different income tax rates for secondary and primary earners, can affect the distribution of market work and housework within the household.

home production, the elasticity of substitution between consumption and leisure becomes almost irrelevant in determining the response of market hours to higher taxes.

In this paper we estimate a discrete choice model of both partners' market labour supply and housework hours. Partners' time allocation choices are modeled as the outcome of maximizing a household utility function with four time uses (his and her market and non-market hours) and household net income as its arguments. The model accounts for (non-)participation in the labour market and housework and incorporates fixed costs of paid work. To approximate continuous hours decisions, each household's choice set is discretized and has 2,401points. The use of a discrete choice specification enables us to incorporate non-linear taxes and (social assistance) benefits.

The model is estimated on data drawn from the 1998-1999 French Time Use Survey. This survey has the advantage of covering a period during which the incomes of French married spouses were taxed jointly and the incomes of cohabiting partners' were taxed separately. Moreover, a time diary was collected for both partners in the household on the same day, which was chosen by the interviewer - in addition to a standard household questionnaire and an individual questionnaire. We observe both partners' market labour supply, housework hours, individual earnings, and household income, as well as the presence and age of children and other individual and household characteristics.

We find positive own net wage elasticities of market work (0.20 for men, 0.55 for women in the baseline specification) and negative own wage elasticities of housework hours (-0.34 for men, -0.36 for women). In absolute terms, an increase in the own wage rate reduces housework hours by less than the increase in own market hours, suggesting that own leisure hours drop as well. An increase in the partner's wage rate reduces own market work hours and increases own housework hours. The elasticities of the husband's market work and housework for the wife's net wage rate are -0.10 and +0.12, respectively; the elasticities of the wife's market work and housework hours for the husband's net wage rate are -0.31 and +0.05, respectively. These cross effects are smaller though than the own-wage effects, as usually found for market work. Own and cross-wage effects on market work are larger for women than for men, which is also a common finding in empirical labour supply studies.

Finally, we simulate the effects of a shift from the current system of joint taxation of married couples' incomes to separate income taxation of married partners.[2] Joint taxation of married couples is mandatory in France. Separate income taxation of married couples is applied in

2 This extends the work of, for example, Steiner and Wrohlich (2004) and Callan, van Soest, and Walsh (2009), who estimated the influence of a similar reform of income taxation for Germany and Ireland, respectively, but only looked at market work of the two partners. It should be noted that the welfare system also has a strong "joint"

most OECD countries. In some countries (for example, the US and Spain), married couples have the option to choose between separate or joint taxation. We find that moving from joint taxation to separate taxation of married spouses' incomes would lead to opposite effects for the husband and the wife: her labour supply would increase while his would fall; and her housework would fall while his would increase. We conclude that replacing joint taxation with separate taxation of married spouses' incomes would increase the wife's participation in paid work by 2.3%-points and her average market hours by 3.7%, while her housework hours would drop by 2.0%. The husband would partly compensate for the changes in the wife's time allocation by increasing his housework hours by 1.3% and reducing his market hours by 0.8%. These effects, though statistically significant, represent only a small step towards balancing market and non-market work of the husband and the wife.

The structure of this chapter is as follows. The model is presented in Section 2.2. Section 2.3 provides an overview of the French income tax system. The data are described in Section 2.4. The estimation results and the simulations are discussed in Section 2.5. Section 2.6 concludes.

2.2 THE DISCRETE CHOICE MODEL

Our model is an extension of the unitary discrete choice model of household labour supply of van Soest (1995).[3] Here we allow individuals in a couple to choose between market work, housework, and leisure while the conventional model allows the individual to choose between market work and everything else and thus, considers housework as "pure" leisure. Hours spent on housework by both spouses enter now directly as arguments of the utility function as individuals choose their hours of market work, housework, and leisure. Therefore, household utility depends on both partners' time allocation and on after-tax household income, which varies with the allocation of hours of market work chosen by the couple, before-tax (or gross) wage rates, and the tax and benefits system. We specify fixed costs of market work and allow for unobserved heterogeneity in partners' preferences. The choice set is discretized and we also include error terms that are specific to each element of the choice set, using a random utility framework.

aspect, since welfare payments are means-tested against total household income. Our simulation leaves the nature of the welfare system unchanged.

3 A discrete choice model of labour supply has also been used by, for example, Aaberge et al. (1995, 1999), Hoynes (1996), and Keane and Moffitt (1998). See also Dagsvik (1994) on the theoretical foundation of the usual functional form assumptions in this type of model.

2.2.1 *Theoretical setup and hypotheses*

Formally, let m denote the husband and f the wife, let t_m^l and t_f^l be the leisure hours of husband and wife, t_m^w and t_f^w their labour supplies, and t_m^h and t_f^h their housework hours. Their gross wage rates are denoted by w_m and w_f. The budget constraint (1) gives family income y after taxes and benefits as a function of gross earnings, total household non-labour income Y_0, and the amount of taxes and benefits T,[4] which depends on the various income components, and on household characteristics X:

$$\begin{aligned} y = {} & w_m t_m^w + w_f t_f^w + Y_0 - T(Y_0, w_m t_m^w, w_f t_f^w, X) \\ & - \mathbf{1}\{t_m^w > 0\} FC_m - \mathbf{1}\{t_f^w > 0\} FC_f \end{aligned} \tag{2.1}$$

The final two terms reflect potential fixed costs of market work, separately for each partner. Fixed costs for the male or female partner enter if that partner participates in market work (where $\mathbf{1}\{.\}$ denotes the indicator function). Non-convexities in the budget set due to taxes, benefits, or fixed costs are allowed for.

The household faces two time constraints given by the total hours endowment E (say 24 hours per day) for each partner:

$$\begin{aligned} t_m^l &= E - t_m^w - t_m^h \\ t_f^l &= E - t_f^w - t_f^h \end{aligned} \tag{2.2}$$

The utility maximized by the household is a function of partners' labour supply, housework, leisure and of after tax household income. Because of the two time constraints, we can eliminate hours of market work and write utility as a function V of five arguments:

$$V = V(t_m^l, t_m^h, t_f^l, t_f^h, y) \tag{2.3}$$

Therefore, household production is not modeled explicitly as for example in Apps and Rees (1999), but is incorporated implicitly by allowing the partners' paid and unpaid housework to enter the model through t_m^h and t_f^h: their marginal utilities not only capture the inherent utility difference between paid work and housework, but also the utility that comes from the household product (which increases with t_m^h and t_f^h).[5] Moreover, the fact that market work is eliminated also matters. In particular, the implications for the expected signs of the partial derivatives of V are as follows:

- $\frac{\partial V}{\partial t_m^l} > 0$ if husband's leisure is preferred to husband's paid work, keeping constant the other arguments of V (including husband's housework and after tax family income y).

4 T also captures welfare transfers (see Section 2.3), which can be seen as negative tax payments.

5 This also implies that we cannot analyze the consequences of policy changes such as a change in VAT that affect the prices of goods bought in the market but not the shadow prices of home produced goods.

- $\frac{\partial V}{\partial t_f^l} > 0$ if leisure of the wife is preferred to paid work of the wife, keeping other factors constant.
- $\frac{\partial V}{\partial t_m^h} > 0$ if housework done by the husband is preferred to paid work done by the husband, keeping other arguments of V constant, including and y. If paid and unpaid work hours are equally attractive or unattractive, we expect [Warning: Image ignored] because housework increases household production, while income from paid work (y) is kept constant.
- $\frac{\partial V}{\partial t_f^h} > 0$ if housework done by the wife is preferred to paid work done by the wife, keeping other arguments of V constant.
- $\frac{\partial V}{\partial y} > 0$ if more household income is better, keeping the allocation of hours chosen by the couple (and therefore also the household production) constant.

only the final inequality is needed to ensure that the model is consistent with the underlying theory as it excludes the possibility that utility falls with income -we assume that the household chooses a point on its budget frontier. There is no need to impose any restrictions on the second order derivatives of V, such as quasi-concavity because to estimate the model we do not have to recur to first and second derivatives –we simply need to compare a finite number of utility values. Finally, the model is static and we do not account for savings (Blundell and Walker, 1986), for a two-stage budgeting approach).

2.2.2 *Empirical specification*

To implement the model empirically, we allow partners to choose their time allocation as follows. We consider 7 discrete possible choices for each activity and for each spouse, which results in a discrete choice set for the household of 7*7*7*7 = 2,401 possible choices. For paid work of men and women, the choices are 0, 1.6, 3.2, 4.8, 6.4, 8 and 9.6 hours per weekday (corresponding to 1, 2, . . . , 6 working days per week). For housework, we use different choices for the two partners (because of the large differences in the observed sample distributions of housework hours of partners, see Section 2.3). We specify 0. 1, 2, 3, 4, 5 and 6 hours per weekday for men, and 1, 2.5, 3.5, 4.5, 5.75, 7.5 and 9.5 hours per weekday for women. For each combination of paid and unpaid work hours of the two partners and for given gross wage rates and household non-labour income, we calculated income taxes and welfare transfers (see Section 3) and therefore, after tax income for each point in the choice set. We assume that partners can choose any combination of hours and ignore possible demand side restrictions (see, for example, Aaberge, Colombino, and Strom, 1999, for an extensive and more complete approach to this issue). However,

our baseline model does incorporate fixed costs of paid work which may partly account for some of these rigidities.[6] We use a flexible quadratic objective function:[7]

We use flexible quadratic objective function:[8]

$$V(\mu) = \mu' A \mu + b' \mu \, ; \mu = (t_m^l, t_m^h, t_f^l, t_f^h, y)' \tag{2.4}$$

where A is a symmetric 5*5 matrix of unknown parameters with entries α_{ij} $(i,j=1,\ldots,5)$, and $b=(b_1, \ldots, b_5)'$ is a five-dimensional vector. We assume that $b_1, \ldots, b_4$ are functions of a vector x of observed household characteristics (such as partners' ages, and the numbers of children in several age groups) and of unobserved characteristics using the following specification:[9]

$$b_j = \sum_k \beta_{kj} x_k + \xi_j; j = 1, 2, 3, 4 \, , \tag{2.5}$$

Here the four unobserved heterogeneity components are assumed to be normally distributed with mean zero and arbitrary covariance matrix, independent of the x_k and of other exogenous components of the model, such as the household's non-labour income and the determinants of gross wage rates. To keep the numerical optimization of the likelihood practically feasible, we do not parameterize α_{ij} $(i,j=1,\ldots,5)$ or b_5, but assume they are the same for all households.[10] Fixed costs of paid work are not observed but are modeled as two unknown parameters to be estimated (one for each partner).

Random error terms are added to the utilities of all m=2,401 points in the household's choice set as in Van Soest (1995):

$$\begin{aligned} &V_j = V(t_{mj}^l, t_{fj}^l, t_{mj}^h, t_{fj}^h, y_j) + \varepsilon_j; \; j = 1, 2, ..., m; \\ &\varepsilon_j \sim \text{GEV(I)}; \; j = 1, 2, ..., m; \\ &\varepsilon_1, \varepsilon_2, ..., \varepsilon_m \text{ independent of each other and of everything else} \end{aligned} \tag{2.6}$$

GEV(I) denotes the type I extreme value distribution with cumulative density. It is assumed that each household chooses the option j that maximizes V_j. The assumption on the error terms then implies that the conditional probability that a given combination j is chosen, given observed and unobserved characteristics, wage rates,

6 It may also be argued that each household needs to do a certain amount of housework, particularly if there are children.

7 To simplify the computational burden, the coefficient of income squared is set to zero, following, for example, Van Soest, Das, and Gong (2002).

8 The coefficient of income squared is set to zero. See Van Soest, Das, and Gong (2002), for example, for a discussion of this specification.

9 The index of the household is suppressed.

10 As usual, the utility function is identified up to a monotonic transformation only. This would make it hard to identify the parameters in a more general model.

other household income, and determinants of taxes, is the following (multinomial logit type) probability:[11]

$$\Pr\left(V_j > V_k \,\forall k \neq j|....\right) = \frac{\exp\left(V(t^l_{mj}, t^l_{fj}, t^h_{mj}, t^h_{fj}, y_j)\right)}{\sum_{k=1}^{m} \exp\left(V(t^l_{mk}, t^l_{fk}, t^h_{mk}, t^h_{fk}, y_k)\right)} \tag{2.7}$$

The scale of the utility function is thus fixed by the magnitude of the common variance of the error terms. The errors can be interpreted as unobserved utility components that make specific combinations of hours in the choice set more attractive than others (in line with the random utility concept in the standard multinomial logit model), or as optimization errors (e.g., errors in the household's perception of the alternatives' utilities).

The probabilities in (7) depend upon the values of the unobserved heterogeneity terms. In order to construct the likelihood contribution of a given household, these terms need to be integrated out. The likelihood contribution then becomes:

$$\Pr\left[(t^l_m, t^l_f, t^h_m, t^h_f) = (t^l_{mj}, t^l_{fj}, t^h_{mj}, t^h_{fj})\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Pr(V_j > V_k \,\forall k \neq j|\xi, ...)p(\xi)d\xi \tag{2.8}$$

Here $p(\xi)$ is the density of the vector ξ of unobserved heterogeneity terms.[12] This likelihood expression involves four-dimensional integrals, which can be approximated using simulations, making it straightforward to estimate the model by simulated maximum likelihood; see, e.g., Train (2003).[13]

The likelihood contribution in equation (2.8) assumes gross wage rates are observed and exogenous. In our data, gross wage rates are not always observed for working individuals and never for non-working individuals. Following most of the labour supply literature, we use separate Heckman models for men and women to deal with unobserved wage rates (see Section 2.4). We then replace either all wage rates or only the unobserved wage rates by predictions based upon the Heckman model estimates. In the first approach, our baseline model, wage rates are allowed to be endogenous, and identification requires variables used to predict wages that do not enter as taste shifters in the labour supply model. Following many earlier studies, we use educational dummies for this purpose. In the second approach (a robustness check discussed in Section 2.5.4), we assume that observed wage rates

11 If hours of work are unobserved but we know that they are positive, the sum of the relevant probabilities is taken, so that the missing information is accounted for.

12 The notation here does not make the conditioning on observed variables explicit, for simplicity.

13 We used 100 Halton draws for each household and each unobserved heterogeneity term; see also Section 2.5.4 for a robustness check.

are exogenous (and measured without error).[14] The difference between the results of the two approaches can be seen as a robustness check for making this exogeneity assumption. Both approaches ignore the potential bias due to prediction errors. In principle, this could be avoided by (for example) estimating wage equations jointly with the structural model. This would, however, substantially increase the computational burden because of the multiple dimensions and because it would require going through the tax and benefits algorithm during each iteration of the maximum likelihood estimation process. Moreover, we would not be able to use a larger sample to predict wage rates (including singles etc.). We therefore could not follow this approach.

2.3 TAXES AND WELFARE BENEFITS

Married spouses are subject to joint income taxation - their incomes are added up for income tax purposes. This typically leads to a larger tax rate for the secondary earner (often the wife) than under separate income taxation. The tax revenue from the joint system is therefore likely to be higher, which means that we could lower the effective joint income tax rates while keeping the tax revenue as high as under the separate income taxation.[15] Most OECD countries have moved to a system of individual taxation or allow couples to choose between the two systems. In contrast to married spouses, cohabiting partners' incomes were taxed separately in France at the time of our survey data.[16] Here we model the income tax system for both married and cohabiting partners.[17]

A key feature of the French income tax scheme is the "quotient familial" ("family quotient" q). Total taxable income is divided by q before applying the tax brackets, and then the resulting amount is multiplied by 3 to give the income tax payable by the household. q gives

14 In this model we still impose the same exclusion restrictions, leading to overidentification. An alternative estimation strategy would be to use data from different years before and after a reform of the tax system, such as the 2000 reform changing the tax treatment of unmarried couples. Our cross-section data did not allow for this.

15 In this revenue-neutral joint income tax system, the effective tax rate for most of the primary earners would be lower than under the separate income taxation, however the effective tax rate for most of the secondary earners would still be higher than under the separate income taxation.

16 Only since the introduction of the "Pacte Civil de Solidarité et de concubinage (pacs)" in 1999, unmarried couples can file jointly, after an initial waiting time of three years. Thus, they could not file jointly before 2002.

17 Different treatment of married and cohabiting couples creates potential for selection into marriage (depending on which form of partnership is more accommodating for the work choices preferred by the spouses). We do not model this form of selection explicitly, but we assess its importance in the policy simulations. We perform our policy simulations separately for the cohabiting and the married households, so that we would have two sets of comparable estimates. If these are sufficiently close to each other, we can conclude that the two samples are not fundamentally different and that the tax-reasons are not likely to play a major role for marriage decisions.

weight one to each married spouse, weight 0.5 to the first and second child, and weight one to children of birth order higher than two.[18] Thus, for a married couple with two children, total taxable income is divided by q=1+1+0.5+0.5=3 before applying the tax brackets, and the resulting amount is multiplied by 3 to give the income which is subject to taxation according to the corresponding tax bracket. In contrast, for an unmarried couple with two children, the two partners file income taxes separately, and thus can choose how to report children for tax purposes. If each of them reports one child, the family quotient for each of them will be 1.5. Combined with the progressive income tax brackets (see below), this system implies that keeping household income constant, the tax paid by a married couple may well be lower than that paid by a cohabiting couple. In particular, a married couple in which only one spouse works and earns, say, y* will pay as much income tax as a married couple in which both spouses work and together earn y* (and much less income tax than a cohabiting couple in which only one spouse works and earns y*). It follows that this system may discourage participation of married secondary earners (see, for example, Apps and Rees, 2011).

The 1998 French income tax brackets that applied to total taxable household income are illustrated in Figure 2.1. There were six income brackets with marginal rates increasing from zero to 54%. The base is gross household income, which is already net of payroll taxes or social security contributions (levied on employers and deducted at source, which are roughly proportional to gross wages); these contributions therefore play no role in the calculations.

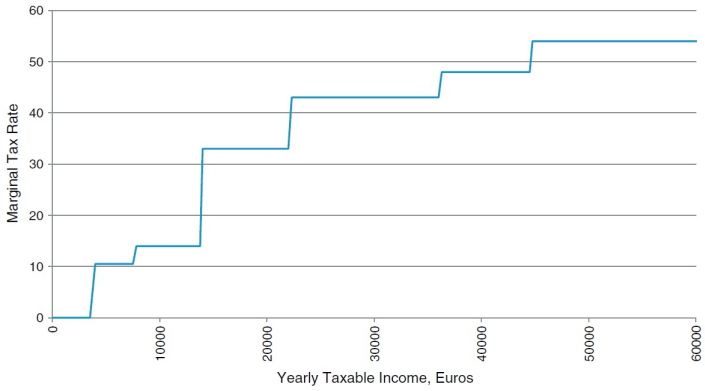


Figure 2.1: Marginal income tax rates for France in 1998

18 Unmarried couples are treated separately with their own family quotients. In an unmarried couple with two children, each partner may, for example, report one child so that each partner's taxable income will be divided by 1.5. Or, if one partner earns more than the other, it may be more beneficial for them if the partner earning more reports both children. In our tax calculations, we always assume that cohabiting partners with children report children for tax purposes so that they minimize the total tax and maximize after tax household income.

1. Standard deductions (on average 28% of total household income[19]) are subtracted from total household income to give 'taxable' household income.

2. Taxable income Y is divided by family quotient, q, which gives the taxable income ratio Y'.

3. The tax rates shown in Figure 2.1 are applied to Y' producing T'.

4. The amount T' is multiplied by q and this gives the income tax payable, T.

5. Low-income households benefit from an additional income tax reduction according to a formula ("la decote") that depends on the income tax payable (T) itself.[20]

According to administrative sources[21] the average (effective) income tax rate for married couples aged less than 60 – the same age cut-off that we use in our sample - is 5.34%, much lower than in most OECD countries, and more than 25% did not pay any income taxes. This is in line with our calculations. For example, a married couple with two children and total annual income of €60,000 has an effective tax rate of approximately 8%, which is low by international standards. It should be noted that unlike in other countries, these French income tax rates do not include social security premiums (which are levied at the source by employers and thus not included in our simulations[22]). Generally, a considerable part of government revenue in France is raised by means of value added tax[23] which we do not model here.

Figures 2.2 and 2.3 show the average tax rate for the household (calculated as the amount of total household tax payable, divided by the total earnings of both partners) as a function of the woman's annual earnings, for various levels of the man's annual earnings for married and cohabiting couples without children (Figure 2.2) and with two children (Figure 2.3). For married couples, the tax rate on each additional euro depends on the earnings of both spouses.

For cohabiting couples, who are subject to individual taxation, the income of the male partner does not matter for the tax rate on the

19 Following, for example, Bourguignon and Magnac (1990), itemized deductions are ignored.

20 If the total income tax payable (T), was less than €508, it was reduced to max (0, 2T-508)., lLow-income cohabiting partners could both benefit from this tax reduction.

21 Enquête Revenus Fiscaux, drawn from administrative income tax files, INSEE, Paris, 1998.

22 The survey collects information on wages net of payroll contributions and gross of income taxes.

23 The amount of revenue levied by means of value added taxes is equal to about 7 per cent of GDP against 10.3 of GDP for income taxes revenue (http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Tax_revenue_statistics). Goods produced within the household such as home cooked meals are not subject to value added tax since the output of household production is hard to measure. In contrast, private goods bought from the market are subject to value added tax.

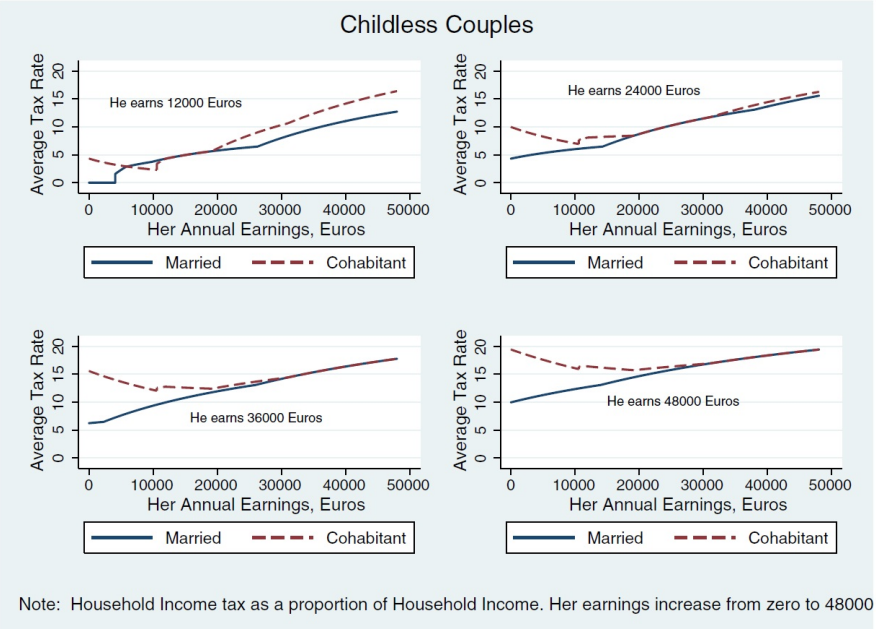


Figure 2.2: Average income tax rates for French childless couples in 1998, keeping men's income fixed

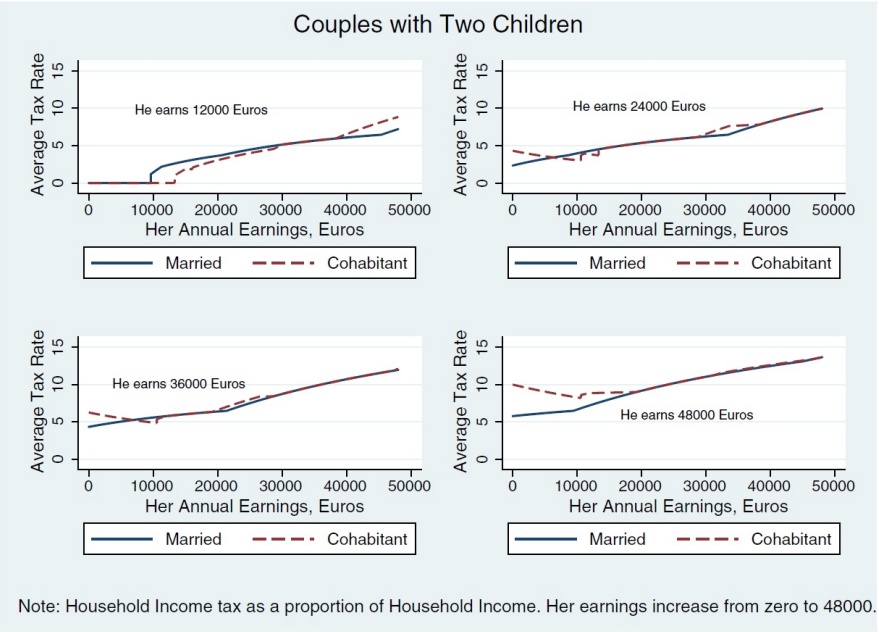


Figure 2.3: Average income tax rates for French couples with two children in 1998, keeping men's income fixed

female's earnings. As a consequence, cohabiting women pay no income tax if their earnings are very low. However, the average household tax rate as a function of her earnings (which is depicted in Figures 2.2 and 2.3), is higher at lower earnings of the female partner in (childless) cohabiting couples than in (childless) married couples (see panels 2, 3 and 4 in Figure 2.2), simply because in married couples the couple's earnings are divided by two (q=2) before applying the tax schedule (see discussion above). If there are children, cohabiting partners can choose who reports them in order to minimize their income tax burden (see also Figure 2.3), and this is the assumption we make in our model, in which we assume that cohabiting couples report their children for tax file purposes so as to minimize the total tax burden. It follows that for various combinations of partners' earnings and family composition, the couple may pay a different income tax for similar total household level depending on marital status (which we take as given here).

In line with the literature on static labour supply models (see, for example, Van Soest, 1995), we do not account for unemployment benefits (which are temporary and depend upon labour market history and involuntary job loss), but we do incorporate the basic social welfare benefits. Their level depends upon the number of children and the benefits are fully means tested on the basis of total household income, regardless of whether partners are married or cohabiting.

We do not explicitly incorporate the costs of child care but control for the presence and ages of children in the model and we include fixed costs of work for both partners.[24]

2.4 DATA

The data for the analysis are drawn from the 1998-99 French Time Use Survey, carried out by the National Statistical offices (INSEE). This survey is a representative sample of more than 8,000 French households with over 20,000 individuals of all ages. Selecting couples, married or unmarried but living together, gave a sample of 5,287 couples with and without children. We further selected couples in which both partners were younger than 60 – the legal minimum retirement age for most workers in France in 1998-99 – and neither spouse was in full-time education, in the military, on disability benefits, or in early retirement.[25] We kept self-employed individuals in the

24 Child care costs of children younger than three vary with the form of child care used by the household but are all tax deductible. Children of age three to six are enrolled in maternal school, which is open 10h a day and free of charge (a symbolic fee is paid for meals, proportional to household income) and almost 100% of French children in this age range are enrolled into maternal school. Older children are enrolled in elementary school which is also open 10h a day and free of charge (a symbolic fee is paid for meals, proportional to household income)

25 We kept housewives as well as men who report that housework is their main occupation (less than ten cases).

sample (whose hours, earnings and total household income were reported in the same way as for employees).

Three questionnaires were collected: a household and an individual questionnaire, and the time use diary. The diary was filled in for one day, chosen by the interviewer, the same day for all household members. About two thirds of the sample filled in the time diary on a week day, and less than a third on a weekend day. We dropped all households who filled in the diary on a weekend day (on which housework is typically not constrained by hours of paid work[26]) or on an atypical day (like a vacation day, a day of a wedding or a funeral, or a sick leave day.), as well as households in which either partner did not fill in the diary. Dropping observations of households who were chosen to complete their time use diaries at the weekend diaries implies that our results refer to partners' time use on week days only. We do not analyze possible (spillover) effects of wages or taxes on house workhousework done in weekends, essentially because we do not observe the same couple on both a week day and a weekend day.

Our final sample for analysis contains 2,141 couples. Table 2.1 shows how many households are deleted from the sample in each of the selection steps described above.

Table 2.1: Sample selection

Selection Criterion	**Households remaining**	**Households dropped**
Original sample size	8186	
Dropping single people	5287	
Dropping couples with one or two spouses older than 59 years	3819	
Keeping in households where both spouses filled in the time diary	3564	245
Dropping spouses that filled in the time diary on an exceptional day	3269	295
Dropping spouses that filled in the time diary on a Saturday or Sunday	2407	862
Dropping people in full-time education or (early)-retirees or doing military service	2141	266

2.4.1 *Sample descriptives, wages and income variables*

Tables 2.2 and 2.3 present descriptive statistics. The average number of dependent children younger than 18 years in the household was slightly over one, implying that 39% of couples in the sample had no children. Only 6% of the sample were not French nationals. Approximately 18% lived in the region of Paris ("Ile-de-France"). Married couples represented 79% of the sample while the remaining 21% were cohabiting. Hourly earnings were computed for respondents who

26 Very few individuals reported any paid work in weekends.

reported continuous (monthly) earnings information, dividing (gross of income tax and net of social security contributions) earnings by usual hours of paid work. The observed average gross wage rates were €9.83 per hour for men and €8.24 for women. Approximately 94% of the men and 70% of the women were engaged in gainful employment at the time of the survey.

Approximately 20% of men and women were self-employed. Average usual hours worked per week were approximately 29 for men and 19 for women, including the zeros for non-workers. Moreover, 360 men and 240 women did not report usual hours, but did report that they were involved in gainful employment. In this case we know that their usual hours are positive and thus, account for this when specifying their likelihood contribution (see Section 2.2 for details).

We predicted wage rates for non-participants as well as for those that did not report continuous wages by estimating a Heckman selection model for men and women separately.[27] See Appendix 2.A for the results. To predict gross (before income tax) hourly wages we use a larger sample than the one used to estimate the model, as we also include individuals that answered the diary on a weekend day or an exceptional day. For estimation of female wages we also include single women in the dataset, assuming that their earnings patterns are similar to those of partnered women. The presence of children and other adults in the household were used to identify the male selection equation from the wage equation.[28] To identify the female selection equation we additionally used marital status dummies, as marital status turned out not to affect female (hourly) wage rates.[29] The presence of children also proved to be insignificant in the wage equations, however the presence of other adults was found to be significantly negative for women. These variables are quite powerful in the participation equation for women but much less so in the equation for men. The selection term is small and insignificant for men, but larger and significantly positive for women, implying that women with unobserved characteristics that make them more productive also have a larger participation probability. The wage equation results are fairly standard, with a mainly increasing quadratic effect of diminishing returns to potential experience and large positive effects of higher education for both men and women. The lack of exogenous source of variation in wages is a drawback of using cross-sectional dataset, which on the other hand is one of the rare surveys to provide detailed

27 Joint wage selection model is in principle feasible, but using it would require discarding part of the female dataset, which we prefer to avoid.

28 Wage rates below half the legal minimum were set to missing (since in some specific jobs it is legal to pay less than the minimum). Wage rate predictions were never below the minimum wage.

29 This is in line with earlier literature that suggests that employers expect all women to marry at some stage and thus apply the same wage 'penalty' to all women, regardless of marital status. Indeed, we found significant wage premiums for married men but not for married women.

information on both partners' time allocation and income. We test for the sensitivity of the estimation results to using observed wages for individuals that reported continuous wages or replacing wages with predicted wages for everyone in the sample.[30]

Table 2.2: Descriptive Statistics

	Husbands		Wives	
Variables	Mean	St dev	Mean	St dev
Age	41.55	9.01	39.25	8.98
Elementary school	0.08	0.28	0.10	0.30
Lower secondary, vocational	0.06	0.24	0.10	0.30
Lower secondary	0.37	0.48	0.28	0.45
Upper secondary vocational	0.06	0.24	0.05	0.22
Upper secondary	0.05	0.22	0.09	0.28
University short degree	0.11	0.32	0.13	0.34
University degree or higher	0.12	0.32	0.10	0.30
French nationality	0.94	0.23	0.95	0.22
Employed	0.94	0.32	0.70	0.47
Self-employed	0.19	0.40	0.21	0.40
Ile-de-France	0.18	0.39		
Regional unemployment rate	11.28	2.35		
Married	0.79	0.41		
Number of children <18 years	1.10	1.12		
Dummy child <3 years	0.16	0.37		
Dummy child 3-5 years	0.15	0.36		
Gross hourly wage predicted	9.77	3.67	6.23	2.55
Gross hourly wage actual	9.85	5.94	8.35	4.92
Usual paid work hours, weekly	29.30	16.57	19.52	17.63
Usual paid work hours, weekly (excluding zeros)	37.94	5.30	32.98	9.01
Paid work (diary), hours - daily	6.97	3.76	4.02	3.93
Paid work, (diary) minutes - daily	418.70	225.51	241.34	235.81
House work, minutes	65.27	85.45	272.49	169.26
Total work, minutes	483.97	196.92	513.84	163.55
"Leisure" (including sleep time and personal care), minutes	956.03	196.92	926.17	163.55

The sample size is 2,141 couples. Hourly wages are gross of taxes. Total work includes paid work, and unpaid housework.

30 The use of observed wages can be desirable, as the observed wages contain more precise information about the incentives faced by individual households. However, their use mandates stronger modeling assumptions - namely, that the part of the wages which is not attributable to observed factors is also uncorrelated with unobservables in the discrete choice model. Such assumption has to be evaluated by comparing the results of models with the observed wages and with the predicted wages.

More than 25% of the sample reported zero non-labour household income (see Table 2.3). Non-labour household income represents approximately 25% of total household income before taxes.[31] The average effective tax rate (the ratio of total household income tax and total household before tax income) is approximately 5.6% of total household before tax income, which is well in line with the administrative data (see also Table 2.1 and Section 2.3). The average effective income tax rate is lower for married couples (5.5% on average) than for cohabiting couples (6.1%).

Table 2.3: Descriptive Statistics: Income and Tax variables

	Q1 (25%)	Q2 (Median)	Q3 (75%)	Mean
Total earnings (€ per year)	12806	21953	32014	23876
Non-labour household income (€ per year)	0	1829	9513	7537
Total household income before tax (€ per year)	21953	28813	37137	31717
Total household income after tax (€ per year)	21108	26783	34426	29187
Total tax burden (€ per year)	0	987	3136	2416
Effective tax rate (%)	1.39	4.49	8.64	5.63

Sample: 2,141 couples. The effective tax rate is defined as the tax amount paid as a proportion of total household income.

2.4.2 *Paid work and housework*

The diary was filled in by each partner on the same week day, which was chosen by the interviewer, spanning 24 hours. Activities were coded in ten minutes slots and approximately 140 possible activities were distinguished by the survey coders. Here, we distinguish the following 'primary'[32] activities

1. Paid work, carried out either at home or at an outside work place.

31 This is before accounting for welfare benefits that are included in our simulation model (see Section 2.3 for details).

32 Respondents were also asked to fill in "secondary" activities which are activities carried out simultaneously, such as cooking while taking care of children. The respondent may report childcare as primary activity and cooking as secondary activity or vice versa. Generally, ignoring secondary activities is likely to underestimate the amount of unpaid work. However, very few respondents in the sample reported some secondary activities, and thus, we resolved to ignore secondary activities. Moreover, if we counted in also time spent on secondary activities, the time budget would not satisfy the 24 hours constraint any longer.

2. Housework, defined to include cleaning, shopping, cooking, doing the laundry, setting and unsetting the table, doing the dishes, doing administrative work for the household as well as any (primary) time spent caring for children.

3. "Leisure" time, defined as any time devoted to leisure (watching television, doing sports, socializing and recreational activities), 'semi-leisure activities' (such as gardening or taking care of pets), as well as personal care and sleeping time.

The distribution of partners' time allocations is illustrated in Table 2.4, which shows that men do the bulk of paid work: the median "husband" in the sample spends approximately 480 minutes (8 hours) on market work, compared to 240 minutes (4 hours) for the median "wife" –denoting the male partner as the husband and the female partner as the wife, for simplicity, even if we included cohabiting couples in the sample. In contrast, women perform most of the housework: with the median "wife" doing 240 minutes of housework against 30 minutes for the median "husband".[33] Interestingly, a comparison of total paid and unpaid work time of men and women in a couple shows that the median "wife" works 10 minutes more than the median "husband" (see also Burda, Hamermesh, and Weil, 2013, on total work load by gender). In the empirical analysis, the time spent on paid work and housework, respectively, by each partner, is rounded to the nearest of the seven discrete point intervals in the choice set (see Section 2.2).

Table 2.4: Time Allocation of Spouses (in minutes on the diary day)

	10%	Q1	Median	Q3	90%
Husband paid work	0	360	480	550	640
Wife paid work	0	0	240	480	520
Husband housework	0	0	30	100	180
Wife housework	70	140	240	390	510
Husband "Total work"	130	420	530	610	680
Wife "Total work"	280	410	540	630	700
Husband "leisure"	740	810	880	970	1170
Wife Total "leisure"	730	790	880	1000	1120

Note: "Total work" time includes paid work and housework. Sample size: 2,141 couples; week day diaries only.

To better grasp within-couple differences in the division of paid and unpaid work, we present the share of the husband's time in the total time devoted by the couple to each activity (see Table 2.5). This shows that the husband provides on average 61% of the paid work done by

33 See also Frazis and Stewart (2012) for a discussion of the limitations of using distributional comparisons.

the couple (and 67% of the median). In contrast, the median husband performs only 12.5% of the couple's housework load. The husband performs on average 45% of the total market and non-market work carried out by the couple (and 47% if we consider the median). To sum up, the wife tends to perform a little more work than the husband (and we have ignored here multi-tasking which is disproportionately done by women, as shown, for example, by Sayer, 2007). Our model will focus on whether this division of time allocations is sensitive to changes in tax rates and other financial incentives.

Table 2.5: The Share of the Husband in Total Spousal Activity Time

	Percentages		
	Mean	St. deviation	Median
Paid work	66.88	30.96	61.07
House work	19.82	22.69	12.5
"Total work"	46.76	15.38	48.78
Leisure	50.08	4.94	50.27

Notes: The shares are calculated only for couples where at least one spouse spends some time on the activity. "Total Work" time includes paid work, housework, and childcare time.

Finally, it should be noted that the number of altenative choices in our empirical specification is rather large compared to the number of households which we observe. This may lead to problems with identification of empirical frequencies of work and housework combinations which are only scarcely chosen (and therefore possibly missing in our data). Potential remedy to this problem is to restrict available choices in the choice set. This can be done either by making the grid of rarely-observed choice combinations coarser, or by discarding certain choices which are, to our conviction, irrelevant for the household decision making. In our study we wanted to avoid these discretional measures, although we admit that there is an ample scope for restricting the choice sets. This can be seen from the Figures 2.A.1 and 2.A.2 in Appendix 2.B where we present frequencies of work and housework intensity levels chosen by men and women in our sample. Particularly for men, the grid of choice combinations contains patches wof combinations with very low frequencies. These parts of the grid could be made coarser without putting excessive restrictions on the flexibility of male decision making.

2.5 RESULTS

In Section 2.5.1 we discuss the estimation results and goodness of fit for the baseline model presented in Section 2.2. Section 2.5.2 presents the wage and income elasticities for this model and Section 2.5.3

discusses the results of the policy simulation of changes in the income tax system. Section 2.5.4 summarizes some robustness checks.

2.5.1 *Parameter estimates and goodness of fit*

We have allowed the parameters of the utility function ($b_1, \ldots, b_4$ in Section 2.2.2) vary with some covariates characterizing the individual and the household (see equation (2.5) in Section 2.2): the age of the individual, marital status, the number of dependent children, and dummies for the presence of children less than three years old or from three to five years old. The systematic part of the utility function therefore contains interactions of leisure and unpaid housework of both partners with these covariates. The parameter estimates of the systematic part of the utility function are given in Table 2.6.

Table 2.6: Estimation Results: Direct Utility functions

Explanatory variables	Coefficient	St. error	
(Husband's leisure)^2	-0.3057	0.0251	**
(Husband's housework)^2	-0.263	0.0171	**
(Wife's leisure)^2	-0.2131	0.0147	**
(Wife's housework)^2	-0.0742	0.0111	**
Income*Husband's leisure	0.0846	0.0089	**
Income*Husband's housework	0.0276	0.005	**
Income*Wife's leisure	0.0564	0.0061	**
Income*Wife's housework	0.0278	0.0038	**
Husband's leisure* Husband's housework	-0.1468	0.0223	**
Husband's leisure* Wife's leisure	-0.0249	0.0068	**
Husband's leisure* Wife's housework	-0.0068	0.0085	
Wife's leisure* Husband's housework	-0.0157	0.0105	
Wife's leisure* Wife's housework	-0.0264	0.006	**
Wife's housework* Husband's housework	-0.0983	0.0118	**
Income	-2.1476	0.4353	**
Husband's leisure	41.7887	7.663	**
Husband's leisure* log age	-17.3115	4.0494	**
Husband's leisure* log age^2	2.4329	0.5536	**
Husband's leisure* married	-0.2621	0.0829	**
Husband's leisure* number children	0.0459	0.0368	
Husband's leisure* any child younger than 3	-0.2048	0.1036	**
Husband's leisure* any child age 3-5 years	0.0341	0.0969	
Husband's housework	15.4829	5.6088	**
Husband's housework* log age	-5.5149	2.8852	*

Husband's housework* log age^2	0.7975	0.3965	**
Husband's housework* married	-0.1988	0.0542	**
Husband's housework* number children	0.114	0.0249	**
Husband's housework* any child younger than 3	0.1786	0.0668	**
Husband's housework* any child age 3-5 years	0.0844	0.0626	
Wife's leisure	52.8154	6.8603	**
Wife's leisure* log age	-25.0188	3.7753	**
Wife's leisure* log age^2	3.4764	0.5264	**
Wife's leisure* married	-0.2381	0.0763	**
Wife's leisure* number children	0.1815	0.0378	**
Wife's leisure* any child younger than 3	-0.1012	0.0876	
Wife's leisure* any child age 3-5 years	0.1924	0.0865	**
Wife's housework	24.4425	4.7226	**
Wife's housework* log age	-11.8946	2.5555	**
Wife's housework* log age^2	1.6968	0.3555	**
Wife's housework* married	-0.0311	0.0489	
Wife's housework* number children	0.2376	0.0243	**
Wife's housework* any child younger than 3	0.2196	0.0536	**
Wife's housework* any child age 3-5 years	0.1558	0.0521	**
Husband's fixed costs of market work	-1.9277	0.1312	**
Wife's fixed costs of market work	-1.3231	0.0945	**

**: significant at two-sided 5% level; *: significant at two-sided 10% level.

The first block of coefficients in Table 2.6 is hard to interpret due to the squares and interactions. Therefore, Table 2.7 presents the average marginal derivatives of the utility function with respect to its five arguments, as well as the fractions of sample observations where the predicted marginal derivative is negative. We find that the objective function increases with the level of household income at every observation in the sample, something that is required for the economic interpretation of the model. For the other marginal utilities, the interpretation in Section 2.2 should be kept in mind. The marginal utility of leisure is negative for almost 27 percent of men and almost 42 percent of women. This indicates that most couples will choose an option with more leisure than paid work if everything else is kept constant (including household income and hours spent on housework). The estimates imply that many respondents would be willing to do some market work for free if there were no fixed costs of work; the substantial fixed costs (cf. Table 2.6) prevent them from doing so.

Table 2.7: Model results: Marginal Utilities

	Average marginal utility	Proportion with negative marginal utility
Income	2.7684	0.0000
Husband's leisure	0.5049	0.2662
Husband's housework	0.0952	0.3092
Wife's leisure	0.3489	0.4199
Wife's housework	0.3546	0.3480

Note: The marginal utilities keep other arguments of the utility function constant. Since paid work is the residual time use category, an increase of husband's leisure implies a fall in husband's paid work, etc.

The marginal utility of housework is positive for 65.2% of women and 69.1% of men, suggesting that, keeping household income constant, non-market work is more attractive than paid work, possibly because of the implied household production output (which is not kept constant; see Section 2.2).

The coefficients on the interactions of exogenous characteristics with the four time amounts in Table 2.6 can be interpreted in a similar way as in van Soest (1995). A positive coefficient on the interaction of a covariate with leisure (of either partner) implies a positive effect of the covariate on the marginal utility of leisure (of that partner) versus paid work, leading to a negative effect on paid hours, ceteris paribus. A positive coefficient on one of the interactions with housework similarly implies a positive effect on the marginal utility of housework versus paid work. For example, the fact that the couple is married rather than cohabiting reduces the marginal utility of the male partners' housework, suggesting that cohabiting men will perform more housework than married men. A plausible explanation is that cohabiting couples are less traditional and have different norms concerning the roles of men and women in the family. As expected, children - and young children in particular - strongly and significantly increase the marginal utilities of both spouses' housework (which includes taking care of children), although the effects are smaller for men than for women.

Table 2.8 gives the estimates of the distribution of the four-dimension vector of random effects in the marginal utilities of leisure and housework time of both partners (cf. Equation (2.5)). The top panel shows that all variances are significantly positive, but their magnitude varies, suggesting that there is more unobserved variation in preferences for leisure (compared to paid work) than in preferences for housework. The bottom panel shows that all correlations are significantly positive, implying, for example, that time use and preferences of both partners are positively correlated, suggesting positive assortative matching.

Table 2.8: Estimation Results: Unobserved Heterogeneity

Covariance Matrix

	Leisure husband	House work husband	Leisure wife	House work wife
Leisure husband	1.4284**			
	(0.1123)			
House work husband	0.3418**	0.1353**		
	(0.0835)	(0.0388)		
Leisure wife	0.7078**	0.3169**	0.7999**	
	(0.0656)	(0.0506)	(0.0649)	
House work wife	0.3144**	0.1788**	0.4683**	0.3051**
	(0.0589)	(0.0312)	(0.0465)	(0.0357)

Correlation Matrix

	Leisure husband	House work husband	Leisure wife	House work wife
Leisure husband	1.0000			
	(0.0000)			
House work husband	0.7764**	1.0000		
	(0.0868)	(0.0000)		
Leisure wife	0.6622**	0.9733**	1.0000	
	(0.0325)	(0.0253)	(0.0000)	
House work wife	0.4754**	0.8905**	0.9483**	1.0000
	(0.0707)	(0.0568)	(0.0174)	(0.0000)

**: significant at two-sided 5% level; *: significant at two-sided 10% level.

Predicted and observed participation rates and mean hours of market and housework are presented in Table 2.9. A comparison of actual and predicted distributions is presented in Figure 2.4. The fit of the distribution of hours spent on housework seems quite good, while that of market work is somewhat less satisfactory, especially for over-time work. It should be mentioned here that incorporating fixed costs helps to fit the participation rates for paid work – models without fixed costs under predict non-participation and over-predict small part-time jobs. The fact that working more than 40 hours is under predicted is probably due to the fact that individuals cannot freely choose to work over time.

Table 2.9: Predicted and Actual Discrete Choice Frequencies

	Husband		Wife	
	Predicted	Actual	Predicted	Actual
		Market work		
0 hours	0.0542	0.0594	0.2938	0.2947
Mean hours	6.8213	6.9106	4.3170	4.6285
		Non-market work		
0 hours	0.4016	0.4340	0.1681	0.1845
Mean hours	1.2943	1.1345	4.6826	4.5636

Note: Hours are per working day (week days only). Market hours are based on usual hours of work per week, divided by five. Actual hours set to missing for those observations reporting no usual market hours but declaring to be in paid employment (360 men and 245 women).

2.5.2 *Wage and income elasticities*

To estimate the sensitivity of the spouses' time allocation decisions to changes in (own or partner's) wage rates and other household income, we have used the estimated model parameters to simulate the distribution of hours of paid and unpaid work of both partners under various scenarios. In each scenario, the discrete distribution (with 2,401 mass points) of time spent by each partner on each activity is simulated for all couples in the sample, accounting for all details of the model such as unobserved heterogeneity and error terms. Unobserved heterogeneity terms are drawn from their estimated distributions, and given the parameters, the unobserved heterogeneity terms, and the budget set in each scenario, the probability distribution over the 2,401 outcomes is calculated for each household in the sample. Based upon these, participation rates in all activities and average hours spent on each activity are computed.

The baseline scenario corresponds to the budget sets used for estimation; this is also the scenario that was used to simulate the predictions in Table 2.9 and Figure 2.4. The other scenarios change the budget sets, either through a change in the net wage rates or through a change in non-labour incomes. For example, raising the net wage rates of all women gives the uncompensated own wage elasticities of paid and unpaid work hours and participation rates for women and cross wage elasticities for men.[34] The net wage elasticities are computed by increasing the net reward for each additional hour of work (by either the male or the female partner) by 1% and comparing the outcomes for these new budget sets with the outcomes of the benchmark simulation. The tax brackets are kept constant under the alternative

34 Changing gross instead of net wage rates gives similar elasticities (somewhat smaller in absolute magnitude).

regime. The net income elasticities are computed by first computing each household's expected income in the benchmark scenario and then raising non-labour income by 1% of this amount for all points in the choice set. Standard errors are obtained by repeating the same simulations for 500 draws of the vector of all estimated parameters from the estimated distribution of the simulated maximum likelihood estimator.

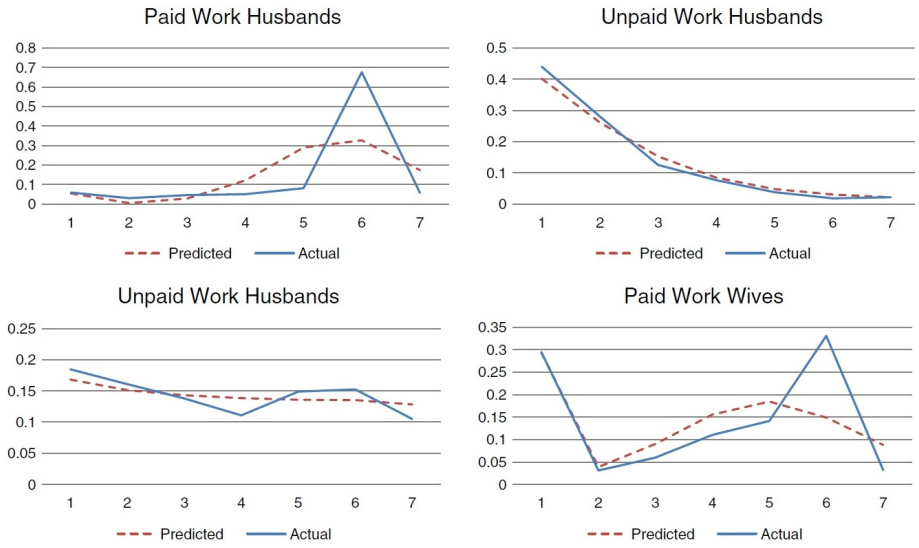


Figure 2.4: Predicted and actual hour frequencies for the (7*7*7*7) discrete choices

Table 2.10 summarizes the results. The estimated female own wage elasticity of market work is 0.55, much larger than the estimate of Bargain, Orsini, and Peichl (2014) but smaller than some of the other elasticities found for France (see Blundell, Bozio, and Laroque, 2013 and the survey in Bargain, Orsini, and Peichl, 2014, Appendix 2.A.1).). The own elasticity is 0.20 for men, which is also larger than in earlier studies. More than half of the estimated responses of the own labour supply to changes in the own wage rate are due to changes in the own participation rate, which is in line with the findings of Bargain, Orsini, and Peichl, 2014.[35]

35 Note that the participation changes are in percentage points; for women (men), the elasticity of participation is about 1.42 (1.05) times as large.

Table 2.10: Own and Cross Wage Elasticities

	Husband		Wife	
	Participation (%-points change)	Average Hours (%-change)	Participation (%-points change)	Average Hours (%-change)
	Market work		**Market work**	
a)Wife's net wage	-0.0087	-0.1039 **	0.2945 **	0.5516 **
1% increase	(0.0085)	(0.0099)	(0.0123)	(0.0371)
b)Husband's net	0.1104 **	0.2025 **	-0.1213 **	-0.3093 **
wage 1% increase	(0.0062)	(0.0184)	(0.0127)	(0.0254)
c)Net family in-	-0.0777 **	-0.1252 **	-0.1628**	-0.2479 **
come 1% increase	(0.0079)	(0.0184)	(0.0203)	(0.0414)
	Non-market work		**Non-market work**	
a)Wife's net wage	0.0412 **	0.1168 **	-0.1734 **	-0.3623 **
1% increase	(0.0079)	(0.0287)	(0.0081)	(0.0225)
b)Husband's net	-0.1940 **	-0.3368 **	0.0344 **	0.0539 *
wage 1% increase	(0.0103)	(0.0564)	(0.0071)	(0.0286)
c)Net family in-	-0.1093 **	-0.3967 **	-0.005	0.0009
come 1% increase	(0.0185)	(0.0568)	(0.0133)	(0.0296)

Notes: **: significant at two-sided 5% level; *: significant at two-sided 10% level. Standard errors in parentheses. Interpretation: In response to an increase of 1% of all women's net wage rates, the women's participation in paid work increases by 0.29%-points and women's hours of paid work increases by 0.55%.

Cross wage elasticities of market hours are negative and smaller in absolute size than the own wage elasticity, but still substantial (and statistically significant): -0.10 for market hours of the husband in response to a change in the wife's wage rate and -0.31 for market hours of the wife in response to a change in the husband's wage rate (again, for simplicity, we denote the female partner as the "wife" and the male partner as the "husband", regardless of the couple's marital status). Also in this case, our estimates are larger than the cross-elasticities found by Bargain, Orsini, and Peichl (2014). For men, most of the cross elasticity is due to changes at the intensive margin; for women, somewhat more than half of the cross elasticity is due to changes at the extensive margin. Estimated income elasticities of market hours are also negative for both partners and equal, respectively, to -0.125 for the husband, and -0.248 for the wife. These are mainly due to responses at the extensive margin.[36] The standard errors indicate that the elasticities of paid work are generally quite precisely determined and statistically significant.

The second panel of the table presents the elasticities of partners' housework to changes in partners' wage rates and net household

36 Income elasticities are not comparable to those in Bargain, Orsini, and Peichl (2014) who only consider changes in capital income and find very small responses.

income. The wife responds to an increase in her wage rate by reducing the time allocated to non-market work - the elasticity is equal to -0.362. In absolute terms, following an upward change in the own wage, the reduction in unpaid work is smaller than the increase of market work, which implies a drop in her leisure hours. Only a small (but statistically significant) part of the reduction in the wife's housework is compensated by more housework been performed by the husband (the cross elasticity of the husband's housework to a change in the wife's wage rate is 0.117), which gives a rather small overall effect as the husband spends little time on housework in the baseline (1.29 hours per weekday). The significantly positive effect of an increase in the wife's wage rate on the husbands' non-market work is in line with earlier findings by Bloemen and Stancanelli (2014), who did not account for income taxation.

The estimated elasticity of the husband's unpaid work to his own wage rate is negative (-0.337) and larger in absolute value than the corresponding elasticity for his paid work. However, because men perform more hours of paid than unpaid work, the overall effect is smaller in absolute terms for housework than for paid work. It follows that an increase in the husband's wage rate leads also to a reduction in his leisure hours. The cross-effect of the husband's wage rate on the wife's unpaid work hours is only marginally significant and quite small (the estimated elasticity is 0.054). In particular, following an increase in his wage rate, his housework drops and hers increases -not enough though to compensate for the reduction in his housework hours, so that the total housework done by the couple falls. Thus, an increase of either the male or the female partner's wage rate reduces the total housework done by the couple, and possibly leads to more outsourcing of household chores.[37]

Finally, the income elasticity of the housework done by the husband is negative and large in absolute value. In contrast, the wife's housework response to a change in non-labour income is virtually zero and insignificant. Thus, total housework falls if other income increases which may suggest perhaps more outsourcing of housework tasks or, possibly, more "multi-tasking" or "leaving housework undone" (see Sayer, 2007, for more insights into all these options).

2.5.3 *Joint versus separate taxation*

Table 2.11 summarizes the effects of a change in the tax system for married couples, from joint taxation (the actual system) to separate taxation (the system in place for cohabiting couples). Cohabiting couples are not included in this simulation for obvious reasons, as nothing changes for them. As anticipated (see Section 2.3), the reform

37 An analysis of outsourcing of housework is given in Stancanelli and Stratton (2014). It is outside the scope of the current paper.

increases participation and hours of market work of married women, and it reduces market work of married men: average hours of paid work fall by 0.75% for the husband while increasing by 3.66% for the wife. In contrast, housework hours increase by 1.28% for the husband and drop by 2.01% for the wife. Thus, these results suggest that a shift from joint to separate taxation would lead to a slightly more balanced distribution of market and non-market work between the spouses. However, these conclusions are driven by few partners that change their time allocation in response to the reform (less than ten per cent of the couples in our sample; results not shown).

Table 2.11: Simulated Effects of Tax Reforms

	Husbands		Wives	
	Participation (%-points change)	Average Hours (%-change)	Participation (%-points change)	Average Hours (%-change)
	Market work		**Market work**	
Separate taxation	-0.1881 *	-0.7513 **	2.3137 *	3.6599 **
for the married	(0. 1209)	(0. 0066)	(1. 3095)	(0. 0213)
Joint taxation for	0.1627 *	1.0413 **	-2.2528 *	-3.5184 **
the cohabiting	(0. 1149)	(0. 0075)	(1. 2848)	(0. 0189)
	Non-market work		**Non-market work**	
Separate taxation	0.6473 *	1.2767 **	-0.8445 **	-2.0147 **
for the married	(0. 3770)	(0. 0203)	(0. 3822)	(0. 0267)
Joint taxation for	-0.7949 *	-1.7559 **	1.1285 *	2.1869 **
the cohabiting	(0. 4618)	(0. 0261)	(0. 5262)	(0. 0259)

Notes: **: significant at two-sided 5% level; *: significant at two-sided 10% level. Standard errors in parentheses. We only consider couples who are affected by the reform (married couples for the first reform, cohabiting couples for the second reform).

Second, we considered cohabiting couples and simulated their responses to a change in the income tax system for cohabiting couples from separate to joint taxation of partners' incomes. As expected, we find opposite patterns than above (see Table 11): cohabiting women are found to reduce their labour supply and increase their housework hours while the opposite is true for cohabiting men. The size of the responses of married and cohabiting partners differ, though, and this may be explained by compositional effects -cohabiting couples are often younger and have fewer children on average than married couples.

2.5.4 *Robustness checks*

Various robustness checks were carried out presented in Table 2.12. We tested the stability of our estimation results by using a new set of Halton draws to estimate the distribution of the random coefficients. Next, we checked the robustness of the estimates to using the observed wages for individuals that reported continuous wages and replaced wages with predicted wages only for observations with missing wage information - this alternative approach implicitly assumes that the errors of the wage equation are independent of the unobservables of the discrete choice model of partners' time allocation.

Table 2.12: Robustness Checks – Elasticities and Taxation Reform Outcomes

Net income elasticities	Baseline specification	Alternative draws	Reported wages	No fixed costs
Market work husband	-0.1252	-0.0418	-0.2053	-0.1813
Non-market work husband	-0.3967	-0.3347	-0.4099	-0.4276
Market work wife	-0.2479	-0.2488	-0.2172	-0.3172
Non-market work wife	0.0009	0.0115	0.027	0.0155
Net wage elasticities (husband's wage)	Baseline specification	Alternative draws	Reported wages	No fixed costs
Market work husband	0.2025	0.2124	0.226	0.2465
Non-market work husband	-0.3368	-0.4087	0.0094	-0.1265
Market work wife	-0.3093	-0.3049	-0.2392	-0.1348
Non-market work wife	0.0539	0.0217	0.0485	0.0178
Net wage elasticities (wive's wage)	Baseline specification	Alternative draws	Reported wages	No fixed costs
Market work husband	-0.1039	-0.0938	-0.0895	-0.0194
Non-market work husband	0.1168	0.105	-0.0723	-0.0549
Market work wife	0.5516	0.5567	0.4556	0.464
Non-market work wife	-0.3623	-0.3191	-0.3749	-0.3597
Separate taxion (married couples only)	Baseline specification	Alternative draws	Reported wages	No fixed costs
Market work husband	-0.0751	-0.0752	-0.0603	-0.0601
Non-market work husband	0.1277	0.1425	-0.0069	0.0261
Market work wife	0.366	0.3614	0.2839	0.2628
Non-market work wife	-0.2015	-0.168	-0.192	-0.1714

Notes: Each column presents one model. The first model is the baseline model discussed in the remainder of Section 2.5. See Tables 10 and 11 for explanations of the elasticities and policy effects.

Net income elasticities	Baseline specification	Part time costs	Fixed & part time costs
Market work husband	-0.1252	-0.0248	-0.0418
Non-market work husband	-0.3967	-0.2984	-0.3347
Market work wife	-0.2479	-0.2683	-0.2488
Non-market work wife	0.0009	0.0117	0.0115
Net Wage elasticities (husband's wage)	Baseline specification	Part time costs	Fixed & part time costs
Market work husband	0.2025	0.1352	0.1258
Non-market work husband	-0.3368	-0.1663	-0.2391
Market work wife	-0.3093	-0.2623	-0.2086
Non-market work wife	0.0539	0.0412	0.0138
Net Wage elasticities (wive's wage)	Baseline specification	Part time costs	Fixed & part time costs
Market work husband	-0.1039	-0.0608	-0.0416
Non-market work husband	0.1168	0.0234	0.0395
Market work wife	0.5516	0.4446	0.3829
Non-market work wife	-0.3623	-0.2701	-0.2499
Separate taxion (married couples only)	Baseline specification	Part time costs	Fixed & part time costs
Market work husband	-0.0751	-0.0423	-0.0363
Non-market work husband	0.1277	0.0437	0.0765
Market work wife	0.366	0.3073	0.2612
Non-market work wife	-0.2015	-0.1514	-0.13

Notes: Each column presents one model. The first model is the baseline model discussed in the remainder of Section 2.5. See Tables 10 and 11 for explanations of the elasticities and policy effects.

Furthermore, we re-estimated the model without allowing for fixed costs of work. Alternatively, we modeled restrictions to the availability of part-time jobs, including and excluding fixed costs of work. Finally, we estimated a simplified version of the model without housework, letting partners choose between leisure and market hours, ignoring housework in the model.

Using a new set of Halton draws (see Column 2 of Table 2.12), some of the estimated elasticities are slightly different but the qualitative conclusions are not affected. Column 3 shows the results if we use actual wage rates for workers whose wage rates are observed (see

discussion in Section 2.2), requiring the additional assumption that errors in the wage equation are independent of the unobservables in the model. The main differences with column 1 are the wage effects on the male partner's housework which seem counterintuitive: a positive (but virtually zero) elasticity for the male's own wage rate and a negative elasticity for the female's wage rate. Accordingly, the policy effect on the husbands' housework is also negative. A possible explanation is that the exogeneity assumption on wages is not satisfied, which would make these estimates inconsistent. The errors in the wage equation are therefore not independent of unobservables both in the first-stage of Heckman model, as well as in the discrete choice model.

Assuming the absence of fixed costs of work (see Column 4 of Table 12), the results are not affected in terms of the direction of the effects but their size differs quite substantially, relative to our favorite specification. Moreover, this specification fits the data worse than our preferred model (see earlier working paper version of the paper that did not account for fixed costs of work). In contrast, simulating restrictions in the availability of part-time jobs (as in Aaberge, Dagsvik, and Strom 1995), including or excluding fixed costs of work, improves the fit of the model (results not shown). Under this scenario, the direction of the effects studies is the same as in our preferred specification but the size of the estimates varies some time quite substantially (see Columns 5 and 6 of Table 12). However, this specification results in more frequent negative marginal utilities of leisure and housework (results not shown) than for our baseline specification, and it is unclear whether this framework would be reasonable to assume here, as while in other countries like, for example, Italy there is a reported lack of part-time jobs, we are not aware of similar issues for France. Therefore, we prefer to retain our main specification. Finally, we assumed that partners only choose between various combinations of paid-work and leisure, ignoring housework (which is then taken as equivalent to leisure), as in most earlier discrete choice models of family labour supply (such as, for example, Callan, van Soest, and Walsh 2009). This simplified model leads to estimated elasticities that have the same sign as those in our preferred model though the size of the effects varies somewhat (see Column 7 of Table 12).

2.6 CONCLUSIONS

We study the impact of income taxation on partners' hours of market work and domestic work in French couples. The theoretical household taxation literature concludes that income taxation is likely to affect not only market labour supply but also housework. However, it is difficult to sign a priori the effect of income taxation on partners' housework. Income taxation is likely to affect labour supply and housework hours in opposite directions because, for instance, downward changes in

the net rewards from work reduce the opportunity cost of housework, making market work less attractive than housework.

There is limited empirical evidence available of the effects of income taxation on housework. Our model extends earlier work on discrete choice family labour supply models by modeling not only partners' market work but also partner's housework. The model accounts for participation as well as hours decisions. The use of a discrete choice specification enables us to incorporate non-linear taxes and welfare benefits in the household budget set. The choice set has 2,401 points for each couple in the sample, since we have allowed for seven discrete paid market-work intervals and seven discrete unpaid-work intervals, for each spouse. Using French time use data to estimate the model, we find that both partners' time allocation decisions are responsive to changes in wage rates, household non-labour income, and the income tax system. In particular, we simulate a change from joint taxation of the incomes of married spouses, which is mandatory in France for married couples, to separate taxation, which was mandatory at the time of the survey for cohabiting couples.

We find that partners' housework responds significantly to changes in the own and the partner's wage rate. The wage elasticities of partners' housework hours are generally smaller in absolute value than those of paid work. We also conclude that replacing joint taxation with separate taxation of married spouses' incomes, would increase the wife's participation in paid work by 2.3%-points and her average market hours by 3.7%, while her housework hours would drop by 2.0%. The husband would partly compensate for the changes in the wife's time allocation by increasing his housework hours by 1.3% and reducing his market hours by 0.8%. These effects, though statistically significant, represent only a small step towards balancing market and non-market work of the husband and the wife. Had we not allowed for housework in the model, we would conclude that the husband's leisure time increases while the wife's leisure time drops following the tax reform.

To sum up, we find sizable wage elasticities of partners' hours of market and non-market work but only small responses of partners' hours to the simulated change in the income tax system, from joint to separate taxation of married spouses' incomes. This may perhaps be due to the small effective income tax paid by French households on average, equal to approximately 5 per cent, on average –social security contributions are levied on the employers' payroll and not accounted for in our model. Having some policy change at hand would enable one to better identify the causal relations at stake, possibly also allowing one to endogenize marital status, which was taken here as given. Future studies should tackle these issues, as well as perhaps model weekend hours (spill-over) effects, which are neglected due to the data availability. We can speculate what would happen had we

observed the weekend time-use patterns for the analyzed families: It may be the case that the couples where both spouses are full-time employed will leave most of their housework for weekends, whereas the couples with one spouse in the household will prefer to spend their weekends enjoying leisure time. This intertemporal substitution of housework would then lead our model to overpredict the substitution patterns between work and housework. Consequently, the wage and tax changes simulated by the model with housework hours would prove to have weaker housework effects than the ones we measure here. Another potential extension is to allow for labor demand constraints - one method to do so would be to generalize our model using the random opportunity specification of Aaberge, Colombino, and Strom (1999). This model allows part-time jobs to differ in their availability from full-time jobs, reflecting the fact that women may be facing more (or less) job opportunities depending on the desired work hours. This extension is pursued as one of the robustness checks in Chapter 3.

2.7 ACKNOWLEDGMENTS

We thank the editor Ian Walker and an anonymous reviewer for many useful comments. We are grateful to the French Agence National de la Recherche (ANR) for financial support. Earlier versions of this paper were presented at a Manheim conference on tax simulation models, an IZA workshop on income taxation, a Nice workshop on the economics of couples, the feminist economics annual conference in Turin 2008, and the International Association for Time Use Research annual conference in Paris in 2010 and at seminars given at San Diego State University, Cergy Pontoise University, Rockwool Foundation Copenhagen, Siena University, and Manheim University. We thank all participants for comments. All remaining errors are ours.

2.A HECKMAN SELECTION MODELS FOR WAGE RATES

Table 2.A.1: Heckman selection model for men

Wage Equation	Coefficient	St. error	
Potential experience	0.0457	0.0038	**
Potential experience squared	-0.0005	0.0001	**
Elementary school	0.0635	0.0359	*
Lower secondary, vocational	0.3110	0.0491	**
Lower secondary	0.2334	0.0376	**
Upper secondary vocational	0.4805	0.0553	**
Upper secondary	0.4998	0.0565	**
University short degree	0.7352	0.0580	**
University degree or higher	1.1217	0.0527	**
Dummy Single	-0.1037	0.0395	**
Dummy Cohabiting	-0.0879	0.0296	**
Constant	2.9003	0.1021	**
Participation Equation			
Potential experience	0.0278	0.0160	*
Potential experience squared	-0.0006	0.0002	**
Elementary school	0.1610	0.1327	
Lower secondary, vocational	0.4884	0.1619	**
Lower secondary	0.3575	0.1074	**
Upper secondary vocational	0.5912	0.1959	**
Upper secondary	0.6287	0.1992	**
University short degree	0.7709	0.1699	**
University degree or higher	0.6137	0.1601	**
Dummy Single	-0.5950	0.1073	**
Dummy Cohabiting	-0.3696	0.0929	**
Dummy child <3 years	-0.1824	0.1051	*
Dummy child 3-5 years	-0.0921	0.1076	
Dummy child 6-10 years	0.0398	0.0947	
Dummy child 11-16 years	0.0179	0.0936	
Number of adults in the household	0.0410	0.0563	
Constant	0.6562	0.2990	**
Inverse Mills Ratio	-0.0224	0.1732	
Implied correlation between error terms	-0.0713		
Standard deviation error wage equation	0.3139		
Observations	2193		

**: significant at two-sided 5% level; *: significant at two-sided 10% level.

The dependent variable in the wage equation is the gross hourly wage rate. Regional controls are included in both equations.

Table 2.A.2: Heckman selection model for women

Wage Equation	Coefficient	St. error	
Potential experience	0.0409	0.0036	**
Potential experience squared	-0.0004	0.0001	**
Elementary school	0.0982	0.0364	**
Lower secondary, vocational	0.2714	0.0390	**
Lower secondary	0.2960	0.0339	**
Upper secondary vocational	0.4291	0.0471	**
Upper secondary	0.5536	0.0414	**
University short degree	0.8458	0.0421	**
University degree or higher	1.2190	0.0447	**
Constant	2.7120	0.0775	**
Participation Equation			
Potential experience	0.0606	0.0119	**
Potential experience squared	-0.0012	0.0002	**
Elementary school	0.3586	0.0927	**
Lower secondary, vocational	0.4606	0.1011	**
Lower secondary	0.4097	0.0833	**
Upper secondary vocational	0.6459	0.1339	**
Upper secondary	0.5527	0.1100	**
University short degree	0.8452	0.1088	**
University degree or higher	0.8370	0.1211	**
Dummy Single	0.4280	0.0720	**
Dummy Cohabiting	0.1811	0.0718	**
Dummy child <3 years	-0.5643	0.0776	**
Dummy child 3-5 years	-0.4890	0.0756	**
Dummy child 6-10 years	-0.3029	0.0626	**
Dummy child 11-16 years	-0.1165	0.0607	*
Number of adults in the household	-0.0655	0.0361	*
Constant	-0.5646	0.2163	**
Inverse Mills Ratio	0.0940	0.0419	**
Implied correlation between error terms	0.2879		
Standard deviation error wage equation	0.3264		
Observations	3406		

**: significant at two-sided 5% level; *: significant at two-sided 10% level.

The dependent variable in the wage equation is the gross hourly wage rate. Regional controls are included in both equations. Dummy Single and Dummy Cohabiting were dropped from the wage equation since they were insignificant.

2.B EMPIRICAL FREQUENCIES OF WORK AND HOUSEWORK COMBINATIONS

The empirical frequencies are expressed as shares of all observed work and housework combinations for the given gender. 1 = 100% of the observations.

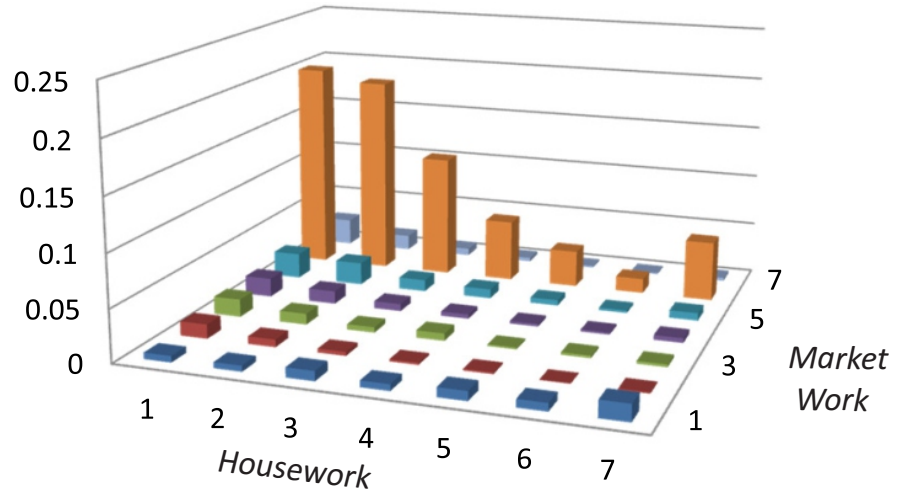


Figure 2.A.1: Empirical Frequencies of Work and Housework Combinations for Men

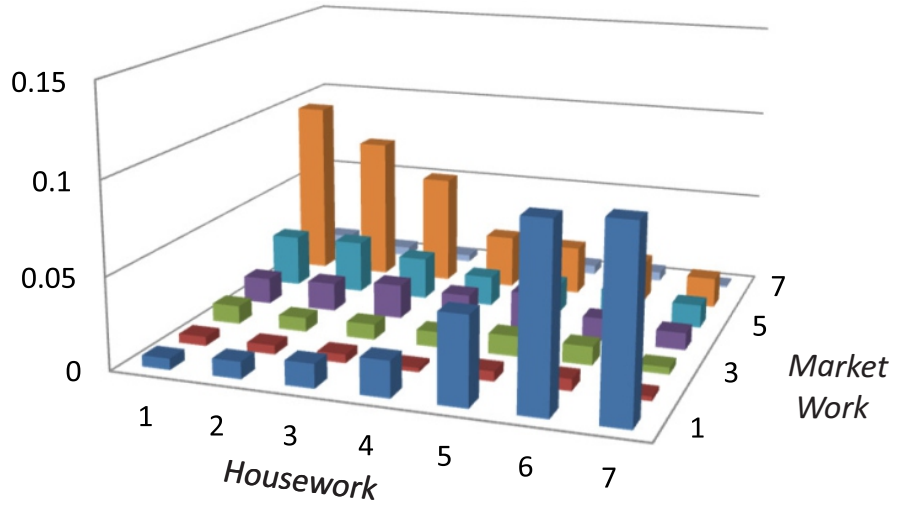


Figure 2.A.2: Empirical Frequencies of Work and Housework Combinations for Women

3

LABOR SUPPLY HETEROGENEITY AND DEMAND FOR CHILD CARE OF MOTHERS WITH YOUNG CHILDREN

This chapter is the reproduction of a paper written with Patricia Apps, Ray Rees and Arthur van Soest, currently being revised for publication in Empirical Economics.

3.1 INTRODUCTION

One of the most striking, and still largely unexplained, facts about female labor supply in the developed countries is its heterogeneity across households, and indeed across countries. In many OECD countries, on average around one third of partnered women work full time in the labor force, one third do various amounts of part time work, and one third work solely in household production. Very little of the aggregate heterogeneity across all households in any one country is explained by wage rate differences and by the number of children present in the household. Moreover, the correlation between female labor supply and fertility across these countries is strongly positive, even though historically, in any one country, there has been an inverse relationship between them.

Some insight is gained by organizing the data in terms of life cycle phases based on the number and age of children in the household. In the pre-children phase, there is very little difference between male and female labor supply distributions. This changes dramatically when children arrive, and this is when the heterogeneity in female labour supply essentially sets in. Though there is a trend of return to the labor force over subsequent phases of the life cycle as the children reach school age and beyond, the basic pattern of heterogeneity persists. Such findings suggest that for the theoretical and empirical analysis of female labor supply it is fruitful to focus on the life cycle phase in which households have young children.

The dramatic change in female labour supply with the birth of the first child reflects the additional work choice created by that event. At least one parent, typically the mother, can choose between working at home providing her own child care or working in the market and buying in care from formal care providers, such as kindergartens and child care centres, or by engaging other care givers, including relatives and friends. The importance of the availability and cost of child care for the labor supply of mothers with young children has been confirmed by theoretical Apps and Rees (2009) and empirical studies, includ-

ing Ribar (1995), Blau (2003), Connelly and Kimmel (2003), Doiron and Kalb (2005), Kalenkoski, Ribar, and Stratton (2005), Kornstad and Thoresen (2007), Baker, Gruber, and Milligan (2008), and Blundell and Shephard (2012). If the mothers are not able to obtain child care from other sources than their own, they will cease working in the market. This lapse in their labor engagement is going to have lasting impact on their later employment decisions. Their human capital depreciates and they become less flexible in their time-use decisions, which lowers their chances of finding a desirable employment opportunity (Apps and Rees, 2009). Accordingly, child care has the potential to be one of the sources of household-level heterogeneity of female labor supply which we observe in the data. That is why we build our analysis of maternal labor supply around the concept of child care, accounting for heterogeneity in attitudes and availability of different modes of child care.

The current paper presents a structural discrete choice model of the time allocation choices of partnered mothers with pre-school aged children. The main advantage of the discrete choice approach is that it can account for the non-convex nature of the household budget sets. Within this model, we analyze the decisions of mothers on hours of market work, time spent on child care and domestic work, and hours of formal child care. The main goal is to assess the sensitivity of choices at the intensive and extensive margin of female labor supply, and to capture underlying substitution patterns between the alternative uses of maternal time.

Similar models are employed by Doiron and Kalb (2005), Kornstad and Thoresen (2007), and Blundell and Shephard (2012). We allow for a more flexible household utility function than previous studies (following Van Soest, 1995 and Kabatek, van Soest, and Stancanelli, 2014 and include both formal child care and maternal care in the utility function. Informal child care is also accounted for. Bought in child care can be incorporated in two ways - either indirectly, subtracting child care costs from disposable household income (Doiron and Kalb, 2005; Kornstad and Thoresen, 2007), or directly, with the hours of bought in child care taken as an additional argument of the utility function (Ribar, 1995; Bernal, 2008). We follow the direct approach, implying that formal child care and maternal care can be imperfect substitutes, with their own effects on household utility.[1]

An important aspect of our empirical model is that we incorporate unobserved heterogeneity in the flexible form of latent classes, following Train (2008) and Pacifico (2012). We thus extend the treatment of unobserved heterogeneity beyond the traditional framework

1 Apart from including formal child care parameters in the utility function, we also subtract the child care costs from the household's disposable income. That way, we account for the income effects of child care utilization.

of random coefficient models,[2] avoiding restrictive assumptions on the distribution of the population parameters of the utility function, which we show has a pronounced effect on estimated labor supply elasticities.[3]

The model is estimated on data drawn from the Household, Income and Labour Dynamics in Australia (HILDA) survey which provides detailed information on time use and child care use and corresponding prices. Simulations based on the estimated parameters show that the time allocations of partnered mothers with pre-school children are highly sensitive to changes in net wages and the cost of child care. A policy simulation also suggests that lowering effective tax rates faced by partnered mothers as second earners, by switching from joint to individual taxation, would lead to a substantial increase in their labor force participation and hours of work.[4]

The paper is organized as follows. In the next section we set out the underlying theoretical model. In Section 3.3 we present the econometric specification that we take to the data. Section 3.4 discusses our data set and Section 3.5 presents parameter estimates. Section 3.6 reports the results of policy simulations. Section 3.7 concludes.

3.2 ECONOMIC MODEL

We construct a one period model of mother's time use and child care decisions during the preschool phase of the life cycle of a two-parent family. The time use decisions of the father, taken to be the "primary " earner, are treated as exogenous.[5] In a one period model, potentially important intertemporal effects, such as the anticipated loss of future human capital and employment possibilities from reducing current labour supply, cannot be incorporated explicitly. A mother may continue to work throughout the preschool phase despite facing a very low net wage or negative net earnings after child care costs, as an investment in her long-term career prospects. We can however

2 Applications using this approach include Ribar (1995), Doiron and Kalb (2005) and Kabatek, van Soest, and Stancanelli (2014).

3 Several studies of female labor supply allow for more flexible treatment of the unobserved heterogeneity (Ribar, 1995; Blau and Hagy, 1998; Tekin, 2007), building on seminal works of Heckman and Singer (1984) and Mroz (1999). The latent class approach can be considered a generalization of these models, allowing for additional flexibility.

4 The policy counterfactual in this simulation is the Australian taxation system which was effective throughout the observed period.

5 By treating father's choices as exogenous, we depart from the assumption of full Pareto efficiency that underlies for example the collective model There is however a growing literature that seeks to relax this assumption, for example that based on non-cooperative rather than cooperative household equilibria. The assumption of exogenous male choices would seem to us to be an acceptable approximation in the light of the results of time use studies showing relatively little variation in male time choices in the early child rearing years, with the vast majority of male primary earners working full time.

partially capture these effects in a reduced form sense, through their impact on the marginal utility of market work *vis á vis* leisure or home child care and domestic work. We also take the number of children in the household as exogenous and therefore do not model fertility decisions. Household $h = 1, 2, ...H$, chooses:

- its consumption of a market good x_{ih}, with $i = 1, 2, .., n$ denoting the individuals within the household;
- the mother's leisure consumption l_{2h};
- consumption of a composite household good, y_h, representing child care and domestic work;
- the mother's time input to the production of the household good, t^y_{2h};
- purchases of the market child care good m^c_h.

Consumption is a composite market good with price 1, the mother's gross wage rate is w_{2h}, and the hourly price of the market child care good is p^c_h, which varies across households.[6] This variation may reflect age of the children in child care, regional price differences, and other factors such as quality of the service. The father's leisure and time allocation to household production are taken to be exogenous and therefore denoted by $\hat{l}_{1h}$ and $\hat{t}^y_{1h}$. Given the time endowment constraint, his market labor supply, $L_{1h} = \hat{L}_{1h}$, is also exogenous.[7] The sum of the parents' gross incomes from market supply, $\sum_i w_{ih} L_{ih}$, is denoted by $I_h(w_{1h}, w_{2h})$. Their utility functions are $u_{ih}(x_{ih}, y_h, l_{ih})$, $i = 1, 2$. The children have utilities $u_{ih}(x_{ih}, y_h)$, $i = 3, .., n$, and are modeled as a household public good.

The household is assumed to maximize a household welfare function, concave in utilities,

$$W_h = \Psi_h(u_{1h}(.), ..., u_{nh}(.); \mathbf{e}_h) \; h = 1, 2, ...H \tag{3.1}$$

where $\mathbf{e}_h$ is a vector of exogenously given "environmental" or "distributional" factors which can be interpreted as determining the household's preferences over the utility profiles of its members.[8] This function is based upon some household choice process which need not be further specified.[9]

6 Every variable or function with subscript h can vary across households. Each of these is therefore in principle a contributor to across-household heterogeneity in choices.

7 The exogeneity of husband's labor supply and income is a strong assumption, however as reported later the data indicate that vast majority of Australian men are working full-time. For that reason, we consider the exogeneity assumption justifiable.

8 In principle, the distributional factors could also include the wage rates, but this will not be allowed for in the empirical model.

9 For a detailed exposition of the economics of this type of household model, see Apps and Rees (2009), Ch 3.

The household's budget constraint can be written as

$$\sum_i x_{ih} + p_h^c m_h^c \leq I_h(w_{1h}, w_{2h}) - T(I_h(w_{1h}, w_{2h}), p_h^c m_h^c; n, ..) \quad (3.2)$$

where $T(.)$ is a tax-benefit function which may contain as arguments demographic variables as well as gross incomes and expenditure on bought-in child care.[10] The technology of the household production of y_h is expressed by the function

$$y_h = g_h(\hat{t}_{1h}^y, t_{2h}^y) \; h = 1, 2, ...H \quad (3.3)$$

and there is a time constraint

$$l_{2h} + t_{2h}^y + L_{2h} = T \quad (3.4)$$

where T is a given time endowment. Because we will be adopting a discrete optimization approach, directly comparing values of the household welfare function at all choice opportunities (see Van Soest, 1995), we do not need to impose conditions of convexity or even differentiability on the function in (3). Thus the household can be thought of as choosing the variables l_{2h}, t_{2h}^y and m_h^c that determine consumptions, market labor supplies and income *via* the constraints (2) - (4) in such a way as to yield a global maximum of the function $\Psi_h(.)$. We can obtain a reduced form of this function by substituting from (2) - (4) into (1) to obtain a utility function that depends on these three choice variables as well as net household income Y. This then forms the basis for the empirical model specification.[11]

3.3 ECONOMETRIC SPECIFICATION

In order to specify a discrete choice model we restrict the values of the three choice variables, the mother's labor supply, L_{2h}, her time allocated to household production, t_{2h}^y, and the hours of bought-in child care, m_h^c, to take one of five possible values which can be characterized as "low", "low-medium", "medium", "high-medium" and "high" according to their observed distributions.[12] The five values of each variable yields a grid of $5^3 = 125$ possible discrete choice points from which the household can choose its optimal allocation. The only restriction we impose on the household-specific choice set is that we exclude alternatives which would imply bought in child care costs exceed family income. This restriction applies mainly to households with the lowest disposable incomes and long hours of formal care.[13]

10 For example there may be tax offsets for expenditure on market child care.

11 This substitution requires additional assumption that the goods which are private enter the utility function additively, reducing them into a composite good.

12 For detailed discussion and applications of the discrete approach adopted here see, for example, Van Soest (1995), Van Soest, Das, and Gong (2002) and Pacifico (2012).

13 In other empirical studies, additional household-specific restrictions are often needed to account for infeasibility of certain choices. Kornstad and Thoresen (2007), for example, constrain choice sets of selected households to account for high degree of rationing in Norwegian day care centers.

Dropping the household subscript, we specify for the purposes of our model the vector $\bar{} = [l_2, t_2^y, m^c, Y]$. The leisure variable, l_2, is the residual of the daily time constraint in (4) with $T = 24$. The mother's household production time, t_2^y, is computed as the sum of hours allocated to child care and to other home production activities, many of which may simultaneously include child care.[14] The leisure time can be then interpreted as free time which is reported not to be spent with children (playing with child is encoded as household production). That being said, since we observe only one activity per diary record, we cannot guarantee that parents do not engage in child care as a secondary activity while enjoying their leisure time.

Net household income, Y, is calculated as gross income net of taxes, family tax benefits and expenditure on child care. Gross income is the sum of each partner's earnings and the family's non-labor income. Since household income does not include the implicit value of household production it does not depend on t_2^y. There are therefore 25 possible values of net household income for each household, corresponding to combinations of the five choices of L_2 and the five choices of m^c.

The mother's gross earnings are calculated as the product of her gross wage and hours of market work. Unobserved wages are predicted by a Heckman selection model (Heckman, 1979), with the exclusion restrictions being number of children in the household and the sum of husband's income and family non-wage income.[15] Expenditure on child care is calculated as the product of a household-specific child care price and the household's choice of formal child care hours. Both characteristics are observed for the families who use formal child care. To account for families who do not use formal child care (and therefore do not report a corresponding price), we follow Connelly (1992) and use a predicted price derived from a Heckman selection model with the exclusion restrictions being number of adults in the household (excluding spouses) and distance from grandparents.[16] Sample selection criteria and regression results for both selection models are presented in the Appendix 3.A.

14 A limitation of the HILDA time use data is that only one activity is reported for each episode. In contrast, the Australian Bureau of Statistics (ABS) Time Use Surveys report a second activity when relevant. Child care is almost always a second activity during housework and related activities.

15 Similar income-based exclusion restrictions are used by, *e.g.,* Blundell, MaCurdy, and Meghir (2007) and Sorensen (1993).

16 The imputation of child care prices should be approached with caution because the observed prices can reflect variation in quality of the service. The quality of child care can be endogenous to the regressors used in the Heckman selection model, and hence can distort reliability of the imputed prices. To address these concerns, we estimate an alternative specification of the model which uses imputed prices of child care for all families in the sample. Relative differences in the predictions made by the original model and the alternative specification can be used to assess whether the endogeneity is likely to play a role here.

Unlike formal child care, the informal child care is not treated as a choice, but it enters utility function as a fixed household characteristic which does not vary with choices made by the household. Empirically, informal child care, because of its low cost, is essentially intra-marginal. Working mothers who cannot obtain enough informal care buy additional care from the market and so respond at the margin to its price and quality. If informal child care were unlimited, households would not use market child care but the constraints on the availability of informal care means that the marginal cost of child care is given by the (quality adjusted) price of market care. Among robustness checks we also estimate an alternative specification of the model which does not maintain that informal child care provision is kept fixed.

3.3.1 *Baseline Model without Unobserved Heterogeneity*

We first present the model without unobserved heterogeneity. We take a reduced form of the household welfare function introduced in the previous section, specified as a flexible quadratic function

$$\Psi(\mu) = \mu' A \mu + \mathbf{b}'\mu \tag{3.5}$$

where $\mathbf{A}$ is a symmetric 4×4 coefficient matrix, and $\mathbf{b}$ is a 4-component vector. The first three components of $\mathbf{b}$, corresponding to the time use variables l_2, t_2^y, m^c, are defined as

$$b_j = \sum_{k=1}^{K} \beta_{kj} X_k, \;\; j = 1, .., 3 \tag{3.6}$$

where the X_k denote respectively a constant term and variables representing observed household characteristics: wife's age; wife's age squared; number of pre-school age children; number of school-age children; and hours of informal child care provided by relatives, friends or the husband. These represent sources of observed heterogeneity. The elements of the matrix $\mathbf{A}$ as well as the component b_4 are assumed the same for all households.[17] Furthermore, in order to prevent marginal utility of income turning negative, we restrict the quadratic coefficient of income to be zero.

The household welfare function in reduced form does not explicitly separate the parameters of the household production function, the utility functions of the household members, or the household process which combines the utilities of the members. This should be kept in mind when interpreting the parameters. For example, the partial derivative of $\Psi(.)$ with respect to t_2^y is the marginal change in household welfare with the other components of $\bar{}$ - l_2, m^c and Y - held

17 This helps to reduce the computational complexity of the problem. Given that the utility function is identified up to a monotonic transformation only, it does not seem overly restrictive.

constant, that is, when an hour of market work is replaced by an hour of work at home without changing income. This captures the (positive) effect of additional home production as well as the potential (positive or negative) effect of a higher or lower preference for home rather than market work, not accounting for the value of home production or the wage for market work. Differences in b_1 across households may therefore reflect either differences in productivity in household production or differences in preferences, or both. Conceptually, these are of course two quite distinct sources of heterogeneity, but they cannot be separately identified in the available data.

We introduce randomness in the value of the household welfare function at each possible choice point (l_2, t_2^y, m^c, Y) by specifying:

$$\Psi_r = \Psi(.) + \varepsilon_r, \ r = 1, 2, .., 125 \tag{3.7}$$

We can rationalize these errors as being errors of optimization or as being due to unobserved alternative specific characteristics that make each alternative more or less attractive than predicted by the systematic part. They can be due to factors that make a specific alternative more (less) attractive because of high (low) productivity or other, possibly preference-related, factors. The ε_r are assumed to be independent of each other and identically distributed and to follow the Type 1 Extreme Value Distribution. This implies that the conditional probability that point r^* is chosen as the optimal point is

$$P[\Psi_{r^*} > \Psi_r, \ \forall r \neq r^* \mid \mu, \mathbf{A}, \mathbf{b}] = \frac{\exp \Psi(\mu_{r^*}, \mathbf{A}, \mathbf{b})}{\sum_{r=1}^{125} \exp \Psi(\mu_r, \mathbf{A}, \mathbf{b})} \tag{3.8}$$

Finally, to guarantee that household welfare always increases with household income (an assumption which is needed for economic interpretation of the estimates) we penalize the likelihood when necessary by adding points inside the budget frontier as additional choices that are never chosen by the household[18].

3.3.2 *Unobserved Heterogeneity*

It is likely that different households within the selected sample of families with young children have different unobserved attributes, for example in human and physical capital, which may impact on home productivity, measured, for example, by child outcomes. There may also be unobserved variation in the quality of market child care. Unobserved heterogeneity, whether in home productivity, in market child care or in preferences, is captured by the specification of error

18 The penalized choices are identical to the standard choices in all respects other than household income, which is lower for penalized choices. Inclusion of such choices in the choice set forces the marginal utility of income to be positive. Negative marginal utility would imply that the penalized choices should be favored by the decision makers, but this is, by construction, never observed.

terms ε_r in the model as interdependent across alternatives. This contrasts with the basic model in which the errors are alternative-specific, which implies independence of irrelevant alternatives.

Several alternative approaches have been developed to allow for unobserved heterogeneity in the context of discrete choice labor supply models. The most prominent one is the parametric random coefficients model (see Van Soest, 1995, or Keane and Moffitt, 1998). This method has been criticized for the restrictive assumptions imposed on the distribution of stochastic terms (see Burda, Harding, and Hausman, 2008; Train, 2008; Pacifico, 2012). The distributions are predominantly assumed to be multivariate normal or log-normal, which implies that the corresponding density of parameter values is unimodal, that is, it has one peak characterizing the most frequent household welfare function. The restrictiveness of the unimodality assumption is well documented in Burda, Harding, and Hausman (2008) who show that the standard random coefficients models perform poorly when the distribution of unobserved heterogeneity has multiple modes. This is not well captured by standard models, rendering the resulting preference ordering too uniform. This issue is of particular importance for our analysis, because previous theoretical work (Apps and Rees, 2009) suggests that multimodal parameter distributions might well be present in the context of female labor supply.

A small body of literature on female labor supply allows for more flexible treatment of unobserved heterogeneity (Bernal, 2008; Blau and Hagy, 1998; Tekin, 2007). These studies draw on Heckman and Singer (1984) and Mroz (1999), using a step function to model the unknown distribution of the key random coefficient. This random coefficient therefore follows a discrete distribution and if enough mass points are allowed for, this distribution is very flexible and can approximate any underlying distribution.

The latent class model can be seen as a tractable generalization of this approach, allowing for flexible discrete distributions of all parameters of the utility function. The underlying assumption is that the population consists of a number of different homogeneous populations (or classes) $K_c, c = 1, ..., C$, characterized by utility functions with parameters $\mathbf{A}_c, \mathbf{b}_c$ (see Train, 2008). The parameterization is class-specific, implying that the probability mass is assigned to the whole set of parameters. This allows individual random coefficients to be correlated, although the correlation structure is not explicitly modeled.

Given the probability $P(h \in K_c)$ that a household $h = 1, ..., H$ is in the class $K_c, c = 1, ..., C$, and writing the probability that point r^* is chosen by this household as

$$P[\Psi_{r^*} > \Psi_r,\ \forall r \neq r^* \mid \mu, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X}] = \frac{\exp \Psi(\mu_{r^*}, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X})}{\sum_{r=1}^{125} \exp \Psi(\mu_r, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X})} \quad (3.9)$$

the unconditional probability that alternative r^* is chosen by household h is

$$\sum_{c=1}^{C} P(h \in K_c) \times P[\Psi_{r^*} > \Psi_r,\ \forall r \neq r^* \mid \mu, \mathbf{A}_c, \mathbf{b}_c, \mathbf{X}]\ ,\ c = 1, ..., C \quad (3.10)$$

The importance of treatment of unobserved heterogeneity is indicative of misspecification of the baseline model. In other words, the true effects of covariates in the baseline model are confounded by variables which are not included in the model. Naturally, there are alternative ways to improve the inference of the model. For example, we can add more covariates into the utility function, or we can estimate the model for isolated subsets of population which can be distinguished by their observable characteristics.[19] Nevertheless, it is unlikely that all the factors influencing the maternal decision making (including child care quality, or cultural norms) would be rendered available to the econometrician. Many of these factors are difficult (or impossible) to observe and therefore they are not included in the conventional datasets. This omission makes the treatment of unobserved heterogeneity a particularly important part of our analysis.

Allowing for multiple latent classes makes the model more difficult to estimate, with the traditional maximum likelihood optimization methods often failing to converge. Train (2008) and Pacifico (2012) show that in such cases we can take advantage of the well-known EM algorithm. This estimation procedure is considerably faster and more stable than the traditional methods, which makes it feasible to estimate flexible models even with a large number of latent classes.

3.4 DATA

The HILDA panel survey provides data on a wide range of variables for a representative sample (17,000 respondents) of the Australian population interviewed annually since the year 2001. Particularly relevant to this study are the detailed data on time use and cost and utilization of formal and informal child care.

Mothers with pre-school aged children represent only a small fraction of each wave of the HILDA survey. To increase sample size we construct a pooled cross-section using the four consecutive waves of HILDA from 2005 to 2008. From each wave we select partnered mothers with pre-school children. We exclude couples in which a partner is disabled, retired, or a full-time student, the husband is unemployed or the family lives in a multi-family household. We also exclude records with incomplete or implausible survey responses (usually on the rele-

19 These methods can be also combined with the methods designed to treat unobserved heterogeneity, as illustrated in Chapter 4, where we estimate separate latent class models for families with children aged 0-3 and families with children aged 4-11.

vant time use variables). The self-employed women are kept in the sample, representing 12% of working mothers.[20]

The final sample contains 1465 records. Descriptive statistics for the dependent variables and the socioeconomic and demographic characteristics entering as independent variables in X are reported in Table 3.1. To enable comparisons by gender, the table also includes descriptive statistics for male wage rates and time use.

Table 3.1: Summary Statistics, Sample of Couples with Preschool Children

Variable	Mean	Std. Dev.	Min.	Max.
Mother's age	32.9	5.6	16	48
Father's age	34.9	6.3	18	58
Marital status (dummy)	0.82	0.4	0	1
Mother's employment status (dummy)	0.55	0.5	0	1
Father's employment status (dummy)	1	0	1	1
Number of children aged 0-4	1.39	0.6	1	4
Number of children aged 5-9	0.47	0.7	0	4
Total number of children	2.01	0.96	1	6
Mother's market work, weekly hours	13.9	15.8	0	80
Father's market work, weekly hours	44.5	11.3	0	128
Mother's h'hold production, weekly hours	71.6	31.2	0	166.8
Father's h'hold production, weekly hours	31.2	17.5	0	120
Mother's leisure, weekly hours	82.3	27.6	0	160
Father's leisure, weekly hours	91.9	18.01	10	150.5
Formal child care, weekly hours	8.3	12.8	0	100
Formal child care price, in AUD, hourly	8.7	3.5	1.583	23.8
Informal child care, weekly hours	6.9	13.6	0	120
Mother's annual earnings in AUD	18514	23570	0	182256
Father's annual earnings in AUD	63373	35736	5136	357216
Annual non-labor family income in AUD	6617	31778	0	683974
Number of observations		1465		

20 This subset of workers is generally excluded from the labor supply analyses, since it is difficult to determine effective wage rates for self-employed workers. The HILDA survey however contains very detailed income section which lists household earnings from a multitude of possible sources, allowing us to approximate earnings even for the self-employed. The inclusion of self-employed may well prove beneficial for the sake of our analysis, because self-employment allows for more flexible working hours, enabling women to combine child care and market work. It should be however noted that since we observe relatively few cases of self-employed women, we do not account for selection into self-employment in the context of our model.

On average, parents of pre-school children are in their early thirties, with the father around two years older than the mother. Only 56 percent of mothers in the sample are employed and, as we would expect, market hours distributions differ dramatically by gender, as shown in Figure 3.1. The result is a gap of over 30 hours per week between average female and male labor supplies. The vast majority of men work full-time (more than 35 hours per week[21]) while women have a distribution of market hours that is relatively uniform apart from a large spike at zero hours. There are 83 mothers who report working more than 18 hours a day for seven days a week.[22] In these cases we scale down the reported hours to satisfy a time constraint of 18 hours of market work and house work per day while retaining the same relative time allocations as in the original data.

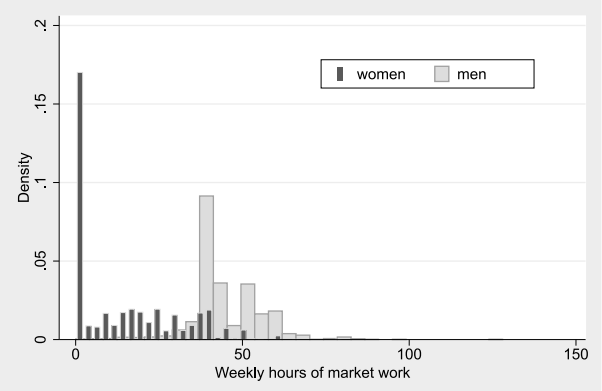


Figure 3.1: Distribution of Weekly Hours of Market Work in Families with Preschool Children

Figure 3.2 compares hours spent on household production activities by gender. As noted above, household production is defined to include the allocation of time to activities involving direct interaction with children, such as "playing with your children", and domestic work, much of which may also involve supervision of children aged 0 - 4. As we would expect, household production hours are higher for females than for males, as shown in Figure 3.3, and their leisure hours[23] are more dispersed, with substantially higher frequencies at the lower levels of weekly leisure time. It is clear that for this group of households with young children, the total work burden is on average

21 “Full-time" employment is defined by the ABS as 35 hours or more per week.

22 The time use data are collected by questionnaire and reported as weekly time uses. Unlike diary data, questionnaire data are typically subject to larger reporting errors, and as a result the sum of individual time allocations to the various activities sometimes fails to satisfy the time constraint.

23 Leisure is computed as the remainder of the daily time endowment after subtracting market work and household production hours, which may be adjusted to satisfy the total time constraint. The 42-hours threshold is following from an assumption that everyone needs at least 6 hours per day for sleep and personal care.

greater for mothers than for fathers.

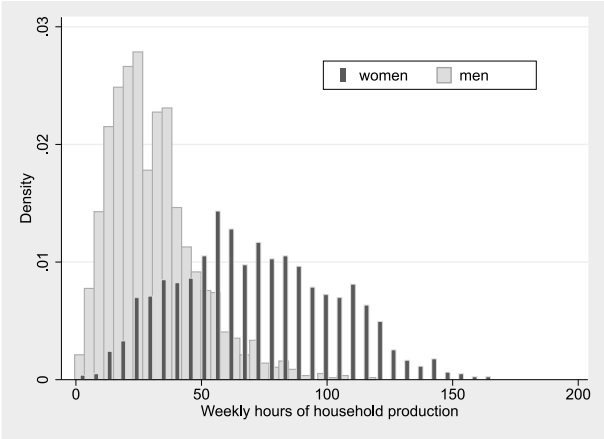


Figure 3.2: Distribution of Weekly Hours of Household Production in Families with Preschool Children

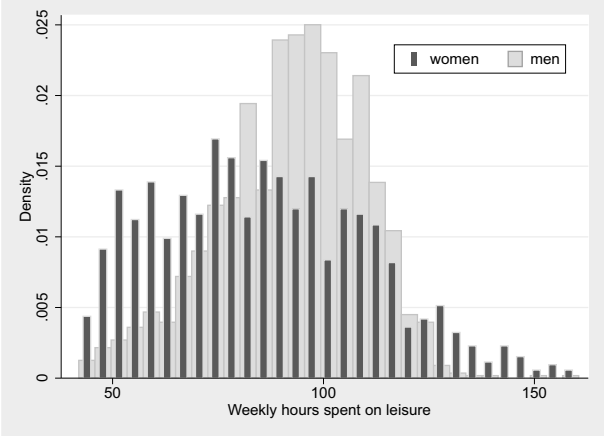


Figure 3.3: Distribution of Weekly Hours Spent on Leisure in Families with Preschool Children

We differentiate between formal child care provided by recognized institutions, such as kindergartens and care centers, and informal care provided by the father, grandparents or other relatives, and friends, for two reasons. First, formal child care differs from informal child care in that it is recognized as incurring costs by the Australian fiscal authorities, and the family is eligible for reimbursement of a considerable part of these costs. Second, the price data on informal care is rather unreliable. The price of formal child care is reported for all children in registered care. In contrast, informal child care is often provided with no charge, or at a price that implies an unobserved subsidy from the carer. The lack of more detailed information about the costs of informal child care makes any effort to impute corresponding prices infeasible. Nevertheless, since the informal child care provision

is assumed to be fixed at the observed levels, we do not need to specify the costs of the service.[24]

Formal care is used by 43% of the families, while the use of informal child care is almost universal (only 9 families report that they used no form of informal child care). The distributions of the weekly hours of child care are presented in Figure 3.4. The profiles for both types of care are relatively similar, although the formal care distribution does not go far above 60 hours per week. This reflects the fact that formal care centers are closed on weekends.

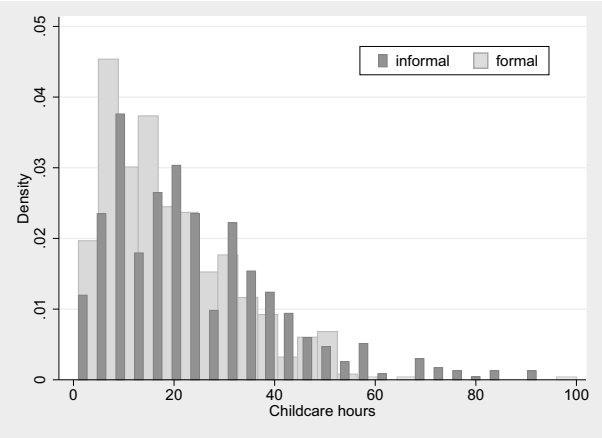


Figure 3.4: Distribution of Weekly Hours of Informal and Formal Child Care, Families with Preschool Children Using Child Care

Annual labor incomes are derived from reported weekly gross salaries from all jobs. The annual non-labor income of the couple is computed as the sum of each partner's business income, investment income, private domestic pensions and overseas pensions.[25] Figure 3.5 presents distributions of male and female labor incomes and household non-labor income. According to these data, around 45% of mothers have zero labor income, while 54% of families in the sample have zero non-wage income. The distribution of non-labor income for the subsample of families with non-negative incomes is skewed towards zero. At the same time several outliers report very large incomes from business and investments.

These income data are used to derive the set of 25 family incomes, net of the taxes and benefits and cost of child care, associated with the discrete time use choices. All incomes are deflated to 2005, the selected base year, using the Australian consumer price index.

24 The costs of informal child care will be reflected only in the informal care parameters included in the utility function. The parameters will however also reflect pure preferences for the informal service, making the separation of informal child care costs infeasible.

25 These pensions are financial transfers which are provided to the male partners prior to retiring officially in the Australian system, while they are still actively working.

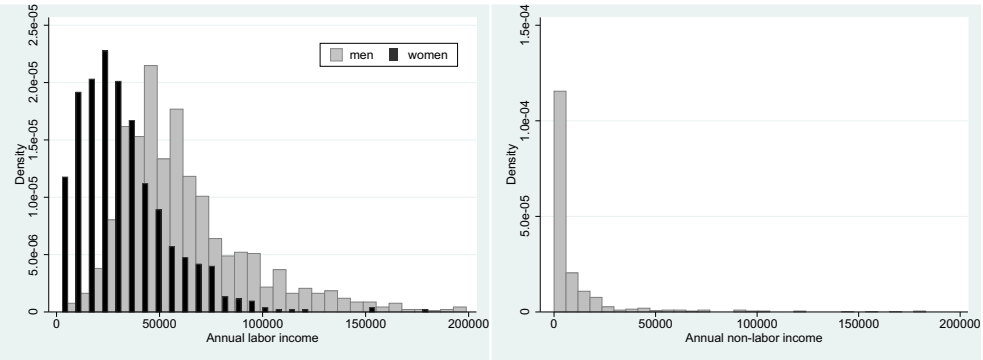


Figure 3.5: Annual Labor and Non-Labor Gross Incomes of Families with Preschool Children, 2008, AUD

Income taxes and Family Benefits

The Australian income tax system is based on individual incomes and the *formal* rate scale is strictly progressive. However strict progressivity is partially lost with the phasing out of a low income tax offset. While tax rates and the offset vary across the four waves of HILDA, the basic structure of the system is essentially the same in each year. For the purpose of illustration, Figure 3.6 plots the profiles of marginal and average tax rates with respect to individual taxable income for the 2007-08 financial year. Details of the rate scale and offset for that year are provided in Appendix 3.B.

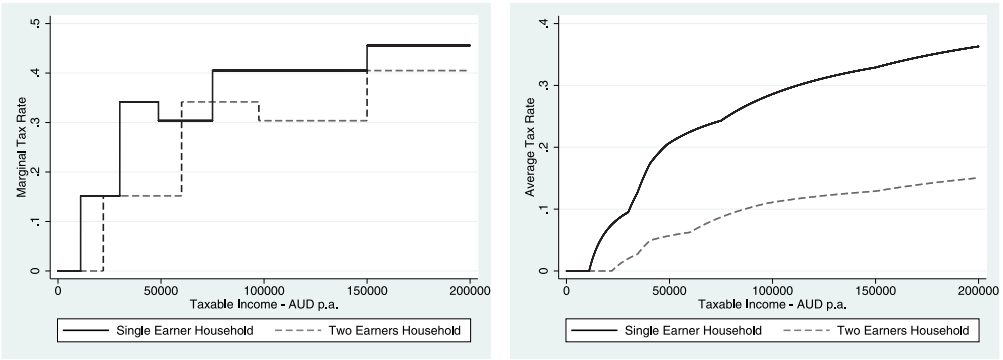


Figure 3.6: Marginal and Average Income Tax Rates, 2007-08, and Annual Incomes, 2008

Because the system is based on individual incomes the marginal rates faced by partners are *independent*. This means that under a progressive rate scale a married mother who decides to switch from untaxed home production to taxed market work faces a lower marginal tax rate, and therefore a lower average tax rate, than her husband if she earns a lower income. In contrast, under a system of joint taxation the tax rates faced by partners are *interdependent*. For the mother who decides to go out to work, the first dollar of her earnings is

taxed at the rate applying to the last dollar of her husband's income. This "tax penalty" on the second earner's income under joint taxation explains why switching from a joint to a purely individual based system stimulates the participation of the second earner (see, for example, Steiner and Whrohlich, 2004).

While the Australia's income tax is based on individual incomes, families are taxed *effectively* under a system of "quasi-joint" taxation. This is due to the withdrawal of child payments at various thresholds defined on joint income under a complex "Family Tax Benefit" system. The *effective* marginal tax rate, obtained by adding the withdrawal rate of payments to the income tax rate, varies widely across the distribution of earnings and can be well above the top rate of the formal rate scale at relatively low incomes levels. This is illustrated in Figure 3.7 for the 2007-08 financial year, for a family with two children aged under 13, with one under 5 years.[26] The figure plots the profiles of effective marginal and average tax rates with respect to the income of the primary earner for two limiting cases: a single income family and a two income family in which both partners earn the same income. For the latter case, the figure plots the effective marginal and average tax rates applying to the income of the second earner.[27] The higher rates on the second income indicate the tax penalty married mothers can face on entering the workforce under a system of joint taxation.

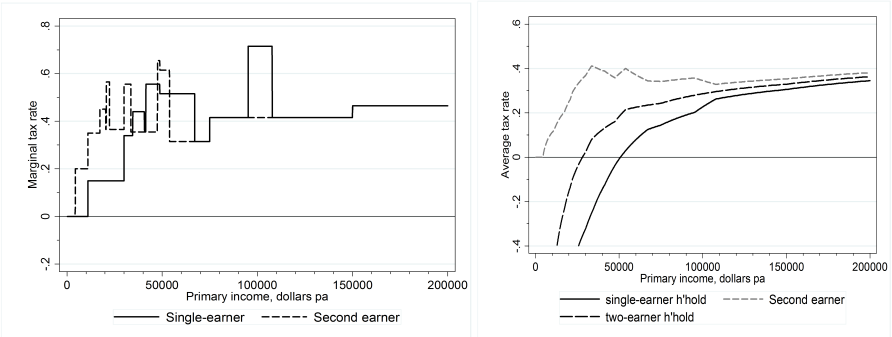


Figure 3.7: Effective Marginal and Average Tax Rates (Including Income Taxes and Family Payments) by Primary Income, 2007-08

The pattern of marginal tax rates in Figure 3.7 implies a budget set with many non-convexities[28] and would make the traditional approach of finding the optimum in the complete budget set infeasible. This makes the discrete approach, approximating the complicated budget frontier with a small finite set of points, particularly useful.

26 The details of the tax rates, family payments and income thresholds on which the figure is based are set out in Appendix 3.B.

27 For a graphical analysis of cases in which the second earner has lower income than the primary earner, see Apps and Rees (2009), Ch 6.

28 For detailed explanation why these non-convexities arise, see Appendix 3.B.

3.5 RESULTS

We first report the results for the baseline homogeneous specification presented in subsection 3.3.1, and then discuss those for the model with unobserved heterogeneity introduced in subsection 3.3.2.

3.5.1 *Baseline Model without Unobserved Heterogeneity*

The estimated parameters of the baseline model are reported in Table 3.2. If the homogeneity assumption were found to be valid, the results would be consistent and more efficient than the latent class model. The coefficients indicate that several of the interaction terms yield intuitively plausible results. An increase in the number of pre-school aged children in the household raises the marginal utility of formal child care, and therefore strengthens the demand for it. On the other hand, an increase in the (assumed exogenous) availability of informal child care weakens it. The same is true for the allocation of time to household production.

The estimated marginal utilities of the choice variables, the components of the vector ¯, are central to our analysis, but their evaluation is more complex than consideration of the simple regression coefficients in isolation, since the marginal utilities depend upon the entire matrix **A** and the vector **b**. They also vary with the household-specific socio-demographic characteristics, **X**, and with the values of the choice variables ¯ themselves. In Table 3.3 we summarize the distribution of the estimated marginal utilities at the observed choices, presenting first their sample averages and second, the proportion of households that have negative marginal utilities. We do this for the full sample as well as for the subsample of households that actually buy formal child care.

As expected, marginal utilities of income, household production and leisure are on average positive, with only a very small fraction of households having a negative value in each case. On the other hand, around 90% of households have a negative marginal utility of formal child care. This is of course not a problem for those households that do not use formal child care, but the last column of the table shows that households that do buy formal care have, on average, negative marginal utilities. This implies that this model is not successful in explaining the use of formal child care from economic arguments. For most households, the use of formal child care can only be predicted with the inclusion of error terms ϵ_r, reflecting optimization errors or unobserved factors that make specific choices more or less attractive.

This counter-intuitive result could be due to unobserved heterogeneity. Our sample contains a large proportion (57%) of households that do not use formal child care. This can be problematic for the homogeneous model if the decision to use formal child is influenced

Table 3.2: Regression Results for the Baseline Homogeneous Model

Matrix A		Vector b	
Income	.199	Formal care*log(age)	-.613
	(.092)**		(3.63)
Formal care	2.23	Formal care*log(age)^2	.130
	(6.20)		(.529)
Household production	4.71	Formal care*married	-.031
	(3.42)		(.057)
Leisure	7.59	Formal care*No. dependent children	-.127
	(3.52)**		(.037)***
Formal care^2	.128	Formal care*children aged 0-4	.329
	(.014)***		(.056)***
Household production^2	.066	Formal care*children aged 5-9	.093
	(.005)***		(.049)*
Leisure^2	.022	Formal care*informal care	-.009
	(.006)***		(.001)***
Income*formal care	.011	H'hold prod.*log(age)	-4.07
	(.004)**		(2.01)**
Income*h'hold prod.	.019	H'hold prod.*log(age)^2	.607
	(.002)***		(.294)**
Income*leisure	.020	H'hold prod.*married	.026
	(.002)***		(.032)
Formal care*h'hold prod.	-.066	H'hold prod.*No. dependent children	-.112
	(.004)***		(.020)***
Formal care*leisure	-.064	H'hold prod.*children aged 0-4	.336
	(.005)***		(.032)***
H'hold prod.*leisure	-.050	H'hold prod.*children aged 5-9	.148
	(.005)***		(.027)***
		H'hold prod.*informal care	-.008
			(.001)***
		Leisure*log(age)	-4.87
			(2.07)**
		Leisure*log(age)^2	.685
			(.305)**
		Leisure*married	.012
			(.034)
		Leisure*No. dependent children	-.053
			(.023)**
		Leisure*children aged 0-4	.231
			(.034)***
		Leisure*children aged 5-9	.114
			(.031)***
		Leisure*informal care	-.009
			(.001)***
n			1465
Log-likelihood			-6076.44

Standard errors in the parentheses, significance levels: 90*, 95**, 99***.

Table 3.3: Average Marginal Utilities of the Main Regressors and Fraction of the Population Sample with Negative Marginal Utilities, Homogeneous Model

	Average Marginal Utility		**Negative Fraction**	
	Full Sample	Child Care Users	Full Sample	Child Care Users
Income	1.06	1.04	0	0
Formal care	-0.69	-0.05	0.83	0.58
H'hold prod.	0.63	0.26	0.17	0.37
Leisure	0.64	0.33	0.24	0.28

by, for example, unobserved differences in home productivity. The model tries to explain this relation in terms of the variables included in the utility function, assigning strong disutility to formal child care. Since the majority of familes do not use formal care, the failure to take account of unobserved heterogeneity forces the common coefficient to be negative. Introducing unobserved heterogeneity may help to solve this problem.

3.5.2 *Latent Class Models*

A key step in the EM estimation procedure is the initial selection of the number of latent classes. This decision involves a trade-off. On the one hand, the higher the number of heterogeneous groups, the better is the fit of the model because we account for unobserved heterogeneity in a more flexible form. On the other hand, more stratified models are bound to be estimated less precisely because the number of unknown parameters rises proportionally to the number of allowed latent classes. The determination of the optimal number of classes is therefore crucial.

Following Train (2008), we compare the models with varying classification choices on the basis of their Schwarz-Bayesian information criteria (BIC)

$$BIC = -2\log(L) + k\log(n) \tag{3.11}$$

where L is the likelihood, k is the number of free parameters in the model and n is the number of observations in our sample. The multiple-class models yield the statistics in Table 3.4. The table shows that the 8-class model attains the lowest BIC, and should therefore be considered as the most reliable specification for further analysis.

In order to examine whether our models actually fit the data, we simulate individual time use allocations using the estimated models and compare the simulated aggregated distributions to their observed

Table 3.4: Bayesian Information Criteria for Multi-Class Models

No. of Classes	Log-Likelihood	BIC
1	-6076.44	12415.31
2	-4921.44	10367.73
3	-4676.13	10139.53
4	-4226.96	9503.61
5	-3963.88	9239.88
6	-3748.64	9071.83
7	-3584.41	9005.79
8	**-3319.71**	**8738.83**
9	-3236.49	8823.87

counterpart. Figure 3.8 presents this comparison both for the baseline model and the model with 8 latent classes.

As expected, the 8-class model replicates the empirical distributions very well, attaining almost identical shares of intensity levels among all three time use choices. The homogeneous model performs much worse and essentially fails to capture the distribution of market work hours. In particular, the model underestimates the proportion of mothers with zero market hours and overestimates the proportion with low hours in part-time work. The distributions of the other two choice variables are replicated well even by the homogeneous model, though the latent class model still provides more precise approximations.

We do not present the regression coefficients for the 8-class model because the class-level stratification makes their interpretation practically impossible. However, one statistic which can be readily interpreted is the fraction of the sample with negative marginal utilities (see Table 3.5).

Table 3.5: Fraction of the Population Sample with Negative Marginal Utilities of the Main Regressors, Model with 8 Latent Classes

	Whole Sample	Child Care Users
Income	0	0
Formal care	0.53	0.31
Household production	0.24	0.41
Leisure	0.23	0.38

The only result which exhibits a substantial change compared to the baseline specification (see Table 3.3) is that for formal child care. The proportion of mothers with disutility from additional formal child care

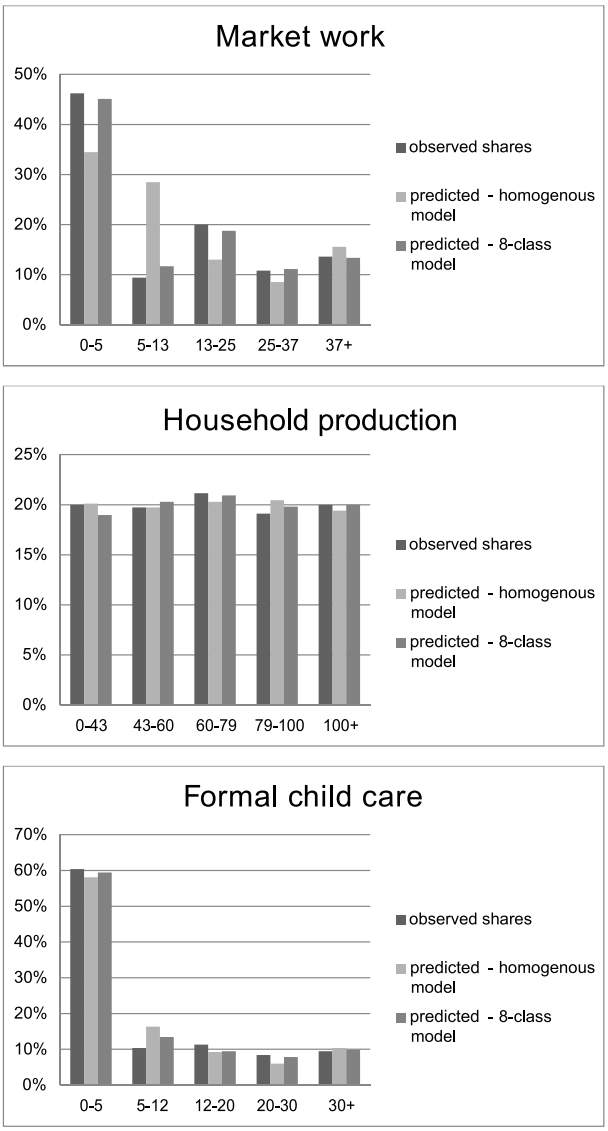


Figure 3.8: Distribution of Time Use Variables into Intensity Levels, Observed and Predicted Shares

drops by 30 percentage points, to 53% in total and to 31% when we restrict the sample to mothers who are using formal child care.[29] This is a considerable improvement over the homogeneous specification.

The relative performance of the models with varying numbers of latent classes is further tested through a series of simulations in the next section. The aim of these simulations is to predict how people respond to selected changes within their economic environment. By predicting (and comparing) the behavioral responses for different model specifications, we can analyze the importance of unobserved heterogeneity and assess the limitations of the homogeneity assumption.

3.6 MICROSIMULATIONS

First, to analyze the sensitivity of choices to wages and prices, we simulate a 10% increase in the net wages of all mothers, and a 10% increase in the net prices of formal child care. Second, we carry out a policy simulation in the spirit of Apps and Rees (2009), building on their critique of joint taxation (as discussed in the previous section). We propose an alternative system of taxes and benefits designed to have a less distortionary effect on female labor supply than the actual system and we estimate its impact on the choices of the type of households we consider.

3.6.1 *Changing Net Wages and Child Care Prices*

The impact of wage and price changes is measured in terms of aggregate elasticities. We compute the percentage changes in total hours of market work, total hours of household production, and total hours of formal child care with respect to changes in the net wage rates of all mothers, or all net child care prices, holding all other variables constant. Changing net rather than gross wages has the advantage of circumventing secondary effects caused by changes in the effective tax rates: increasing net wages by 10% results in 10% higher disposable incomes from the mother's market work across all households.[30]

The resulting income changes are proportional to the net earnings of mothers so that those working earn more while non-participants retain their original disposable incomes. Since the 10% increase in the wage makes participation more attractive, we can expect an increase both in market hours of employed mothers and in the labor market

29 The latter is obtained by taking weighted means over all classes, where the weights are the class probabilities given the observed choice.

30 We decided to present net elasticities here, because they are more reflective of people's attitudes. Gross elasticities render the estimates of labor flexibility specific to the tax system in place (which is often subject to change), making it difficult to compare our results across different studies - especially if the other studies focus on different countries. Interested reader can find the comparison of gross and net elasticities in Appendix 3.C.

participation rate. Similarly, an increase in net child care prices results in an income reduction that is proportional to the cost of bought-in child care, and we can expect that it leads to a reduction in bought in hours and in the fraction of the sample using formal child care.

We compute aggregate elasticities as the ratios of percentage changes in the relevant time or care use to the percentage changes in the wages and prices (where the latter are 10 percent, by construction). This is done as follows. We first derive the benchmark time use allocations (using the same wages and prices that are used for estimation) by averaging individual choice probabilities predicted by the model and using these to compute the average hours of each activity.

A similar procedure is applied to calculate the average hours of activities after the wage or price increase. The only difference is that the choice probabilities are derived using adjusted disposable income for each alternative in the choice set. This changes the utility values for some of the choice alternatives but not for others, and, as a consequence, changes the probabilities of all choices. Using the new probabilities we recompute average hours. Finally, we compute the percentage deviations in the new averages compared to the benchmark. The elasticities for models with a varying number of classes are provided in Table 3.6.[31]

Table 3.6: Elasticities of Time Use Allocations with Respect to Changes in Net Wages and Net Child Care Prices

Mothers' Net Wage Increased by 10%									
No. of classes	1	2	3	4	5	6	7	8	9
Formal care hrs.	1.01	0.58	0.58	0.58	0.59	0.50	0.51	0.42	0.49
	(0.051)**	(0.042)***	(0.096)***	(0.101) ***	(0.070)***	(0.078)***	(0.095)***	(0.065)***	(0.097)***
Market work hrs.	1.35	0.77	0.65	0.48	0.65	0.62	0.52	0.43	0.68
	(0.043)***	(0.034)***	(0.084)***	(0.088)***	(0.067)***	(0.126)***	(0.134)***	(0.086)**	(0.149)***
H'hold prod. hrs.	-0.23	-0.10	-.10	-0.06	-0.10	-0.07	-0.10	-0.08	-0.09
	(0.012)***	(0.011)***	(0.020)***	(0.016)***	(0.018)***	(0.015)***	(0.023)***	(0.015)***	(0.023)***
Net Child Care Price Increased by 10%									
No. of classes	1	2	3	4	5	6	7	8	9
Formal care hrs.	-0.51	-0.51	-0.48	-0.52	-0.47	-0.45	-0.50	-0.42	-0.38
	(0.017)***	(0.02)***	(0.031)***	(0.043)***	(0.033)***	(0.079)***	(0.06)***	(0.081)***	(0.06)***
Market work hrs.	-0.17	-0.11	-0.15	-0.10	-0.12	-0.10	-0.11	-0.08	-0.11
	(0.01)***	(0.011)***	(0.018)***	(0.026)***	(0.02)***	(0.023)***	(0.027)**	(0.021)***	(0.027)***
H'hold prod. hrs.	0.03	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.01
	(0.003)***	(0.004)***	(0.006)***	(0.005)**	(0.007)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***

Standard errors in the parentheses, significance levels: 90*, 95**, 99***.

The first panel gives the responses to the increase in all mothers' net wage rates. The first thing to note is the large difference between the homogeneous (one class) model and the models with unobserved

31 Standard errors on the elasticities were computed through 199 Monte Carlo simulations, recomputing the percentage changes with simulated sets of parameters determining **A** and **b**. These parameters were drawn from the estimated (multivariate normal) distribution of the ML estimates. See Ruud (1991).

heterogeneity (two or more latent classes), demonstrating the importance of controlling for unobserved heterogeneity. When we allow for unobserved heterogeneity, the predicted responses fall substantially, and remain relatively stable among models with different numbers of classes.

Standard errors tend to increase as we allow for more classes, making some of the effects less significant for heavily stratified models. This reflects the fact that these models are more flexible and therefore require more data for accurate estimation. Nevertheless, there are also cases where standard errors fall as we move to the more stratified models. We attribute this effect to the increased goodness of fit of the latter specifications.

Given the results of the BIC selection procedure discussed in the previous section, our discussion of simulation outcomes now focuses on the 8-class model. The net wage increase leads to time use shifts that correspond to intuition. A 10% increase of all net wage rates results in a (significant) 4.3% rise in average working hours, implying a positive uncompensated own labor supply elasticity of 0.43 for this group of mothers with young children. This is well in line with the large literature on female labor supply. The positive substitution effect (the price of leisure increases) dominates the negative income effect. Moreover, the 10% wage increase leads to a (significant) 4.2% increase in hours of formal child care (a "cross" elasticity of 0.42). First, the higher demand on time due to increasing hours of market work leads to substitution of own child care for bought-in child care. Second, higher earnings lead to higher family income, increasing the demand for formal child care if this is a normal good.

The elasticity of time allocated to household production is significantly negative, at -0.08. The negative sign implies that higher wages lead mothers to work less in the household. However, the actual change in home production hours is not large enough to compensate for the increase market hours, implying that mothers also reduce their leisure in order to do more market work.[32]

Turning to the impact of the rise in child care prices in the second panel, it is not surprising that the highest elasticity is that of formal child care itself. With a 10% rise in child care prices, the demand for formal child care falls significantly, by 4.2%. This in turn causes mothers to work less in the market - market hours drop significantly, by 0.8%, as they have to substitute their own time for bought-in services.[33] Accordingly, the hours of household production increase

32 In absolute terms, the wage increase induces the average mother to spend about 0.55 hours per week more on market work, 0.40 hours less on household production, and 0.15 hours less on leisure.

33 This elasticity is well in line with that of Gong and Breunig (2011) but smaller than that in Gong and Breunig (2012).

by 0.2%, replacing almost all of the forgone time formerly spent on market work.[34]

3.6.2 *Simulation of a Tax and Benefit Reform*

As discussed in section 3.4, the phasing out of family benefits on household income creates high effective marginal tax rates for many mothers as secondary earners. To investigate the impact of these high rates on their labor supply and participation, and also on the demand for formal child care, we simulate the effects of switching to an individual based income tax with universal payments. The reform replaces the marginal rate scale depicted in Figure 3.7 with one that applies to individual taxable incomes, as illustrated in Figure 3.6. To fund the increase in benefit payments we increase proportionally (by 26.76 percent) all marginal tax rates of the system in Figure 3.6 to achieve a reform which is ex ante (that is, before behavioral responses) revenue neutral.[35] An alternative to this reform would be to discard family tax benefits, and decrease the income taxes to compensate families for the forgone benefit.

Figure 3.9 shows graphically the differences in the net tax positions of households resulting from the reform (assuming no behavioral responses). The differentials are ordered by the corresponding pre-reform net household incomes, so that we can see how the shift in the tax burden varies with household income. Since the reform is ex ante revenue neutral, the changes for all families in the sample sum to zero. The figure shows that when benefits are universal families with average joint incomes gain. It is important to keep in mind that this is due to gains for relatively low to average two earner families who previously lost the joint income tested benefits. The proportionally higher marginal tax rates shift the tax burden towards the higher income groups, in effect shifting the burden from lower wage two earner families to those with higher wage rates.

Table 3.7 summarizes the simulated changes in time allocations and hours of formal child care in response to the reform. As in the previous simulations, we observe a large discrepancy between the changes predicted by the homogeneous model and those predicted by the models with more than one latent class, with the results of the latter proving relatively stable across different specifications. Again

34 On average mothers spend about 0.10 hours per week longer on household production and reduce their market work by 0.10 hours and bought-in child care by 0.05 hours.

35 The reform is based on the population with preschool children. However, in the context of the Australian family tax system for two-parent families with dependent children, it can be viewed as one that needs to be considered for the full-population. The extremely large child payments increase further for older children, extending the problem of high effective marginal tax rates on secondary earners across the full-population of families with dependent children.

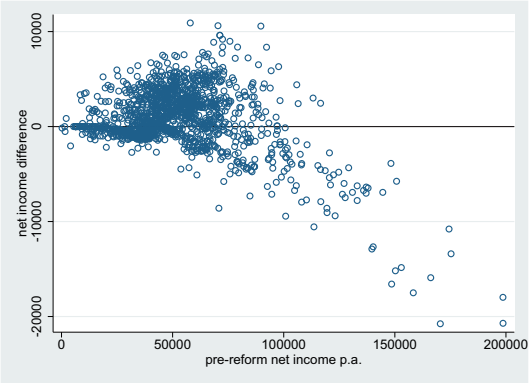


Figure 3.9: Post-Reform Differences in the Net Tax Positions of Families, Ordered by Pre-Reform Net Household Incomes

this implies that accounting for unobserved heterogeneity is important not only to improve the fit of the model but also from a substantive point of view.

Table 3.7: Percentage Changes in Time Allocations after FTB Reform

No. of classes	1	2	3	4	5	6	7	8	9
Formal care hrs.	3.90% (0.243)***	2.77% (0.234)***	2.77% (0.648)***	3.52% (0.634)***	3.78% (0.54)***	1.57% (0.422)***	2.24% (0.511)**	1.75% (0.38)***	2.74% (0.564)***
Market work hrs.	6.49% (0.337)***	4.43% (0.262)***	3.54% (0.627)***	4.16% (0.616) ***	3.96% (0.62)***	2.99% (0.83)***	2.77% (0.885)**	3.11% (0.684)***	3.41% (0.915)***
H'hold prod. hrs.	-1.11% (0.091)***	-0.6% (0.076)***	-0.57% (0.155)***	-0.63% (0.113)***	-0.39% (0.144)***	-0.38% (0.136)***	-0.66% (0.191)**	-0.63% (0.133)***	-0.61% (0.152)***

Standard errors in the parentheses, significance levels: 90*, 95**, 99***.

We again focus on the outcomes for the 8-class specification. We observe that the reform would lead to a 3.11% increase in average hours of work (about 0.43 hours per week, using the average hours in Table 3.1), a 1.75% increase in average hours of formal child care (0.15 hours per week), and a 0.63% decrease in the average hours of home production (about 0.45 hours per week). All these effects are statistically significant. On average, the positive effect on market work of not phasing out family benefits is more important than the negative effect due to the increase in the marginal tax rates. Market hours of work therefore increase and, as a consequence, hours of home production fall and demand for formal child care increases.

Heterogeneity of the Behavioral Responses

The behavioral effects induced by the reform appear to be highly heterogeneous across population groups and latent classes. Closer analysis of our results reveals positive effects at the extensive margin of female market labor supply, with the predicted labor market participation rate rising by 4.4% (to about 58%). On the other hand, these

effects are mitigated by responses at the intensive margin, with some employed mothers choosing to work fewer hours under the reform. Average hours of market work (conditional on being employed) fall by 1.3%, with individual responses showing considerable variation. In fact, expected hours of market work increase for 69% of all women in the sample. The 1.3% decline in the aggregate work at the intensive margin is driven by the response of mothers on higher wages and in full-time employment. Clearly these women are not among the prospective beneficiaries of the reform because the extra family benefit payments are not sufficient to compensate them for their higher tax burdens. Facing lower net incomes they substitute away from market work towards non-market time uses.

Such behavioral heterogeneity is crucial for successful targeting of policy reforms, as it helps to identify the potential impact on different subsamples of the population. It is also interesting from the perspective of economic modeling, as we can compare the relative performance of homogeneous and latent-class models. In order to do so, we split the sample into two groups according to actual employment status, and compute the elasticities separately for the two groups, using both the homogeneous model and the latent class model with eight classes. Using the homogeneous specification, the effects prove to be almost identical for both groups, as this model captures only a small part of the differences in productivity and preferences between the groups (the "observed heterogeneity" part captured by the covariates in the model). According to the eight class model results, the simulated increase of aggregate working hours is much stronger for non-employed mothers, with the absolute increase of market work hours being 28% larger. As for the change of formal child care hours, the non-employed mothers exhibit a rather modest increase in absolute terms (70% lower than employed mothers), but in relative terms their bought-in child care rises more than for employed women (the initial level of formal child care utilization is substantially lower for non-employed mothers).

The failure to capture heterogeneity in responses of the homogeneous model is further illustrated by the fact that this model cannot replicate observed differences in reported time use allocations between the two groups, overestimating work and formal child care allocations of non-employed mothers and underestimating them for employed mothers. On the other hand, the 8-class model produces almost identical time use patterns as observed in the data. For these reasons, it is hard to maintain that the homogeneous model would be able to provide reliable predictions of the responses to proposed policy changes.

Net Fiscal Effect of the Reform

We also analyze the net revenue effect of the reform taking account for behavioral changes predicted by our 8-class model. Changes in time allocations can affect government revenue through two distinct channels: by increasing (reducing) their hours of work mothers are also increasing (reducing) income tax revenues, and by buying in longer (shorter) hours of formal child care, child care benefits rise (decline).

The key result in this context is that the government marginally improves its net fiscal position. Income tax revenue from mothers rises by only 0.5%, which seems low compared to the 3.1% increase in aggregate working hours. The reason is the heterogeneity in responses discussed above: mothers with higher wages tend to reduce their hours of market work, and the progressive nature of the income tax system makes the fall in tax revenues from this group relatively large, substantially offsetting the additional revenue from low and middle income households. More specifically, mothers who increase their market hours (69% of the sample) are predicted to pay an average of \$151 more in annual income taxes (a 5.6% increase) whereas those who reduce their hours reduce their income tax liabilities by a predicted average of \$261 per annum (a 2.3% reduction). The net result is an aggregate increase of the income taxes by \$25 per household (which translates into the aforementioned 0.5% revenue gain).

The situation is very similar for the child care benefits, which increase only slightly in aggregate: on average, a household gets an additional \$2 (0.1% of the initial payments), which is small compared with the 1.75% change of formal child care hours. Analogously to the income tax effects, this outcome reflects the heterogeneity in behavioral responses.[36] Combining the two effects, we estimate that on average households will contribute an additional \$23 to government tax revenue, an increase that represents 0.2% of their original contribution.

3.6.3 *Robustness Checks*

In order to assess the stability of our results, we run a series of sensitivity checks, altering the econometric specification of our model in the following ways. First, to achieve a more flexible specification, we divide the time use variables into a finer grid (6^3) of discrete points, allowing a greater degree of choice in household decision making. Second, we experiment with the composition of time use variables,

36 In our sample, 65% of mothers are predicted to increase their hours of formal child care (typically the mothers who increase their hours of market work). Once again, a key role is played by mothers with higher wages who reduce their market work hours and also their utilization of formal care. The relatively large fall in their claimed benefits is sufficient to offset most of the rise in benefit claims by other households.

reducing the mother's household production decision to a single maternal child care choice.[37] Third extension augments the model by a fixed disutility of work which is estimated as an additional parameter of the utility function.[38] Fourth, in order to account for potential dependence of child care prices on the quality of the service, and for misreporting in the individual household accounts, we estimate a model with imputed wages and child care prices for everyone (instead of just for the households where wages or prices are not observed). Fifth extension investigates plausibility of the assumption that post-reform informal child care use remains fixed at the pre-reform levels. We do so by estimating a model which does not allow for interactions of choice variables with informal childcare. Therefore, this model does not assume that the provision of informal childcare remains unchanged, and the potential effects of substitution between informal and formal childcare are contained within the formal child care coefficients included in the utility function. Sixth robustness check generalizes our model using the random opportunity specification of Aaberge, Colombino, and Strom (1999). This model allows part-time jobs to differ in their availability from full-time jobs, reflecting the fact that women may be facing more (or less) job opportunities depending on their desired work hours. Unlike Aaberge, Colombino, and Strom (1999), we do not allow for availability of jobs to depend on the corresponding wages, as this adjustment proves cumbersome in our model specification. Last extension evaluates the maintained assumption that women's leisure is consumed privately, yielding the same utility irrespective of the leisure choices of their partners. To check its plausibility, we augment the utility function specification with an interaction term for male and female leisure time. The male leisure time is treated as given for the sake of the model. The interaction term allows for preference for shared spousal leisure time, so that women can attribute higher utility to leisure when their partners are enjoying it as well. The next step in loosening the assumptions imposed on father's time use would be to allow him to respond to the choices made by his partner, altering his own labor engagement and own child care provision. Here we do not pursue this extension since a host of empirical literature (see Doiron and Kalb, 2005 or Apps and

37 This allows us to examine direct substitution effects between maternal and formal child care.

38 There is no clear consensus with respect to which form the working indicator should take on. Blundell, Gosling, Ichimura, and Meghir (2007) put the employment dummy into the budget constraint, so that it represents fixed monetary costs of working. Donald and Hamermesh (2009) interact the dummy with time use variables entering the utility function, referring to the corresponding parameters as shifters of time use efficiency. We choose to add the employment dummy into the individual utility function in a non-interacted form, which allows us to model fixed disutility from work without substantially increasing the computational burden.

Rees, 2009) showed that such responses are unlikely to occur in the Australian context.[39]

Table 3.8: Robustness Check - Elasticities and Reform Responses Derived by Alternative Model Specifications with 8 Latent Classes

Mothers' Net Wage Increased by 10%

	Original Model	6 Brackets	Maternal Care	Fixed Disutility
Formal care hrs.	0.42 (0.065)***	0.51	0.43	0.42
Market work hrs.	0.43 (0.086)**	0.61	0.57	0.57
H'hold prod. hrs.	-0.08 (0.015)***	-0.12	-0.07	-0.09

	Imputed	No Inf. Care	Random Opportunity	Interacted Leisure
Formal care hrs.	0.36	0.35	0.51	0.55
Market work hrs.	0.47	0.51	0.52	0.76
H'hold prod. hrs.	-0.06	-0.10	-0.14	-0.11

Net Child Care Prices Increased by 10%

	Original Model	6 Brackets	Maternal Care	Fixed Disutility
Formal care hrs.	-0.42 (0.081)***	-0.45	-0.44	-0.35
Market work hrs.	-0.08 (0.021)***	-0.11	-0.09	-0.11
H'hold prod. hrs.	0.02 (0.008)***	0.02	0.01	0.01

	Imputed	No Inf. Care	Random Opportunity	Interacted Leisure
Formal care hrs.	-0.40	-0.42	-0.43	-0.39
Market work hrs.	-0.10	-0.07	-0.09	-0.11
H'hold prod. hrs.	0.01	0.02	0.02	0.02

FTB Reform - Percentage Changes

	Original Model	6 Brackets	Maternal Care	Fixed Disutility
Formal care hrs.	1.75% (0.380)***	1.94%	2.29%	2.68%
Market work hrs.	3.11% (0.684)***	3.18%	3.01%	4.26%
H'hold prod. hrs.	-0.61% (0.133)***	-0.77%	-0.37%	-0.96%

	Imputed	No Inf. Care	Random Opportunity	Interacted Leisure
Formal care hrs.	1.92%	1.14%	2.58%	2.18%
Market work hrs.	3.25%	3.45%	4.23%	3.73%
H'hold prod. hrs.	-0.58%	-0.83%	-0.87%	-0.76%

Standard errors in the parentheses, significance levels: 90*, 95**, 99***.

39 Nevertheless, in other countries the father's behavior may be of considerable importance - if father is likely to lower his labor supply and increase his child care provision in response to his partner's increased work engagement, then our model would be misspecified. The model would not be able to account for in-house substitution of child care sources, assigning too much importance to the arrangement of formal child care for newly employed mothers. Accounting for flexibility of father's choices would then drive down mother's wage elasticity of formal child care use. Mother's own wage elasticity of labor supply would increase because the woman would have another source of child care at her disposal, making her work decision more flexible.

Table 3.8 shows that changes in the econometric specification induce changes in the exact values of the elasticities, but their relative sizes and signs remain similar to those in the original model. Most of the values remain within the 95% confidence interval of the corresponding baseline elasticities. This finding is particularly important in the context of child care quality concerns, as it suggests that the differences in service quality reflected by variation in observed prices are unlikely to distort our estimates.

The stability of the elasticities is also interesting in the context of the model containing maternal child care decisions, as it suggests that changes in the hours of home production are proportional, irrespective of the distinction between child care-related and other household activities. Women who engage in the labor market will therefore work less in the household, delegating part of their chores either to the husband or buying in the services from the market.

We also check the validity of standard errors corresponding to the measured elasticities without changing the specification of the model itself, but calculating standard errors in a robust way, controlling for general heteroskedasticity and household-specific clustering (considering the estimates as pseudo maximum likelihood estimates). In both cases the newly derived standard errors preserve the significance levels attained by the benchmark approach, suggesting that the heteroskedasticity or clustering is not likely to distort our results.

3.7 CONCLUSIONS

In this paper we have analyzed the time allocation decisions of mothers with pre-school children, with emphasis on the influence of a non-convex tax and benefit system on labor supply, household production and the use of formal child care. We have focused on incorporating unobserved heterogeneity, originating possibly from differences in productivities and preferences. Our findings show this plays a dominant role in analyzing the mothers' decisions in our data. Our results cast strong doubts on the usefulness of the homogeneous model with no unobserved heterogeneity. The parameters fail to capture the true effects of factors driving household decision making, and hence simulations based on the baseline homogeneous model specification give misleading results.

To control for unobserved heterogeneity, we estimated a series of latent-class models, among which the 8-class model was found to perform best, balancing goodness of fit against parsimony. To assess the responsiveness to changes in the family tax system and in child care prices, we conducted several simulations based upon our estimated models, increasing net wages of mothers or net child care prices, and altering the joint-income structure of the existing tax and benefit system in the third reform.

The simulations show that mothers are responsive to changes in wages as well as changes in child care prices. The results suggests that market work and formal child care tend to be complements, and respond significantly to wage and price changes. The results also indicate that significant changes in labor supply and child care demand can remain unidentified when the unobserved heterogeneity is not accounted for, since the homogeneous model leads to significantly distorted female labor supply elasticities.

In the third simulation we show that the phasing-out of family benefits on the basis of joint income increases marginal tax rates on the incomes of mothers as second earners, with a negative impact on their labor supply. The tax system can be made more favorable for mothers by switching to a fully individual based system. In such a setting, women are predicted to increase their labour supply and use of formal child care. The net effect of these responses is to raise additional tax revenue which could be used to lower tax rates and therefore achieve efficiency gains. The gains from the reforms we have simulated arise from changing the structure of effective marginal tax rates under the Australian "quasi-joint" family tax system. Running a similar policy simulation on data for countries with full joint taxation may yield considerably stronger behavioral responses.

A number of improvements and extensions are of course possible. First, our analysis would benefit from exploiting the panel structure of the HILDA data set, controlling for time-stable individual effects. Secondly, although we consider the current method of treating unobserved heterogeneity to perform well, it could be worthwhile to assess the stability of our results by using alternative ways of controlling for unobserved heterogeneity, such as the random coefficient mixed logit model, or the approaches utilizing Bayesian nonparametric methods.

3.8 ACKNOWLEDGMENTS

This research was supported under the Australian Research Council's Discovery Project funding scheme (DP1094021). The authors thank seminar participants at Université Laval, University of Lausanne, University of Mannheim, Netherlands Bureau for Economic Policy Analysis (CPB), Tilburg University, the ESPE Annual Conference, the NATSEM International Child Care, Work, and Family Workshop, the Netspar Pension Day, and the Netspar International Workshop for their valuable comments.

3.A HECKMAN SELECTION MODELS

Here we present details of the Heckman selection models used to predict missing wages and child care prices. The wage regressions are estimated on a sample of partnered women aged 55 years and under, with those reported as full-time students or disabled excluded. The exclusion restrictions used in the participation equation are non-wage income and number of children.

Table 3.A.1: Mothers' Wage Estimation, Heckman Selection Model. The Dependent Variable Is Gross Hourly Wage.

	Participation Equation	Wage Equation
Constant	-.027	3.984
	(.307)	(7.380)
Married	-.048	1.057
	(.039)	(.925)
Urbanization index	.100	1.086
	(.028)***	(.730)
Non-English ethnicity	-.287	-4.351
	(.046)***	(1.176)***
Mother's age	.033	.265
	(.015)**	(.366)
Mother's age squared	-.0009	-.006
	(.0002)***	(.005)
Mother's tenure	.032	.404
	(.004)***	(.113)***
Mother's tenure squared	.0005	.0001
	(.0001)***	(.003)
Other household income (log)	-.100	
	(.012)***	
No. of children aged 0-4	-.574	
	(.024)***	
No. of children aged 5-9	-.226	
	(.025)***	
No. of children aged 10-14	-.149	
	(.026)***	
No. of children aged 15-18	.019	
	(.032)	
Inverse Mills Ratio		13.121
		(1.994)***
Obs.	9324	9324

Standard errors in the parentheses, significance levels: 90*, 95**, 99***. Additional controls include regional, educational and yearly dummies.

Both of these exclusion restrictions are standard in the literature[40], however irrespective of their prevalence, it is beneficial to discuss their limitations. Traditionally, there has been a lack of robust exclusion restrictions for selection models of female wages, since most of the variables which influence women's labor participation do influence their wages as well. Our two restrictions may prove susceptible to this

40 Among others, they are used in Connelly, 1992, Kornstad and Thoresen, 2007, or Doiron and Kalb, 2005

critique as well: The non-wage income is likely to influence wages if men match with women of similar earnings potential. The number of children can have also non-trivial effects on wages, either through selection into fertility, or if childbearing is causing women to lose human capital. Some of these issues are addressed in Chapter 5 where we use an alternative approach for modeling female wages.

The child care price regressions are estimated on a sample limited to mothers with pre-school children. This subsample is larger than the sample used for estimating our discrete choice model because the time-use data in HILDA was collected for a randomized subsample only. The exclusion restrictions in the child care participation equation are number of adults in the household (excluding the spouses), and residential distance from grandparents (the base group represents families without grandparents).

Table 3.A.2: Child Care Price Estimation, Heckman Selection Model. The Dependent Variable Is Gross Hourly Price of Formal Child Care.

Wage equation	Participation Equation	Price Equation
Const.	-2.645	8.090
	(.968)***	(6.187)
Mother's gross hourly wage	.003	-.008
	(.0009)***	(.008)
Other household income (log)	-.010	.049
	(.030)	(.121)
Married	-.050	.409
	(.094)	(.414)
Urbanization index	.096	.354
	(.068)	(.288)
Non-English ethnicity	-.197	-.195
	(.125)	(.586)
Mother's age	.069	.001
	(.055)	(.273)
Mother's age squared	-.0003	-.0002
	(.0008)	(.004)
No. of children aged 0-4	.174	-.276
	(.059)***	(.294)
No. of children aged 5-9	-.084	-.130
	(.051)*	(.229)
No. of children aged 10-14	-.274	-.981
	(.087)***	(.471)**
No. of children aged 15-18	-.277	-1.683
	(.143)*	(.652)***

No. of other adults in the household	.015	
	(.131)	
Distance to grandparents: Same household	-.730	
	(.376)*	
Distance to grandparents: Less than 1 km	-.408	
	(.147)***	
Distance to grandparents: 1 to 4 kms	-.243	
	(.144)*	
Distance to grandparents: 5 to 9 kms	-.091	
	(.135)	
Distance to grandparents: 10 to 19 kms	-.086	
	(.138)	
Distance to grandparents: 20 to 49 kms	-.297	
	(.153)*	
Distance to grandparents: 50 to 99 kms	-.025	
	(.210)	
Distance to grandparents: 100 to 499 kms	-.068	
	(.166)	
Distance to grandparents: 500 kms or more	-.030	
	(.154)	
Distance to grandparents: Overseas	-.413	
	(.218)*	
Inverse Mills Ratio		.932
		(1.498)
Obs.	1725	1725

Standard errors in the parentheses, significance levels: 90*, 95**, 99***. Additional controls include regional, educational and yearly dummies.

3.B AUSTRALIAN FAMILY INCOME TAXES AND CHILD CARE SUBSIDIES

Net household income is calculated as gross income net of tax liabilities and family payments. We compute tax liabilities under the Personal Income Tax (PIT), Low Income Tax Offset (LITO), Medicare Levy (ML),[41] and family payments under Family Tax Benefit Part A (FTB-A) and Family Tax Benefit Part B (FTB-B). The calculation of the net price of formal child care takes account of the two main subsidies for child care, Child Care Benefit (CCB) and the Child Care Rebate (CCR). Each of these component of the overall system is described below. We also report details of the tax-benefit system used to construct Figures 3.6 and 3.7 for a family with two children under 13, with one under 5, in the 2007-08 financial year.

Personal Income Tax and LITO

The 2007-08 marginal rate scale of the PIT begins with a zero rated threshold of $6,000, followed by rates of 15%, 30% and 40% up to an income of $150,000, and thereafter a top rate of 45%. The LITO in the same financial year provided a tax offset of $750, phased out at 4

41 Despite its title, the ML is entirely an income tax. It is not tied to funding any aspect of the health system.

cents in the dollar on individual incomes above $30,000. The resulting *effective* rate scale was therefore a zero rated threshold of $11,000 and a higher rate of 34 cents in the dollar on incomes from $30,001 to $48,750, as depicted in Figure 3.6.

Medicare Levy

A Medicare Levy (ML) applies at a rate of 1.5% of income, with exemptions defined on family income and varying with the number of children. In 2007-08 the family income limit for a full reduction for a two-parent family was $29,207, plus $2,682 for each dependent child. The exemption was withdrawn at a rate of 8.5 cents in the dollar above this limit, with the effect of raising the marginal rate above that limit to 44 cents in the dollar. Thus the ML introduces a further nonconvexity in the *effective* rate scale and also shifts the tax base towards joint income.

FTB-A and FTB-B

FTB-A provides a payments for each dependent child. The size of the payment varies with the age of the child. The "Maximum Rate" of FTB-A in 2007-08 for a child under 13 years was $4,460.30. This maximum payment was withdrawn at 20 cents in the dollar on a family income over $41,318 up to the "Base Rate" of $1,890.70 per annum. The Base Rate was withdrawn at 30 cents in the dollar at a higher family income threshold that depends on the number of dependent children. For a family with two dependent children, the income threshold for the Base Rate was $95,192.

FTB-B provides an annual payment of $3,584.30 for a family with a child under 5 years. The payment was withdrawn at a rate of 20 cents in the dollar on a second income above $4,380. It can therefore be classified as a "gender based tax" (see Alesina, Ichino and Karabarbounis, 2011) with, paradoxically, the higher rate applying to the income of the mother as second earner.

Child Care Benefit and Child Care Rebate

Child Care Benefit depends (among other things) on the ages of children, number of children, type of child care and the hours of child care used. The benefit is phased out with rising family income according to the age of the child and the number of children receiving child care.

The Child Care Rebate reimburses families for their claimed child care expenses. It can cover up to 50% of the net child care expenses (that is, after subtracting CCB). The CCR rate is not income-tested, but it has an upper cap on the amount of expenses which can be reimbursed. For the year 2008, this cap was $4,354 per year.

3.C COMPARISON OF GROSS AND NET ELASTICITIES

Table 3.A.3: Elasticities of Time Use Allocations with Respect to Changes in Wages and Child Care Prices, Partnered Mothers with Pre-school children

Mothers' Net Wage Increased by 10%		
No. of classes	1 class	8 classes
Formal care hours	1.01 (0.051)***	0.42 (0.065)***
Market work hours	1.35 (0.043)***	0.43 (0.086)***
Household production hours	-0.23 (0.012)***	-0.08 (0.015)***

Mothers' Gross Wage Increased by 10%		
No. of classes	1 class	8 classes
Formal care hours	0.70 (0.031)***	0.25 (0.069)***
Market work hours	1.02 (0.031)***	0.31 (0.099)***
Household production hours	-0.17 (0.008)***	-0.05 (0.012)***

Net Child Care Price Increased by 10%		
No. of classes	1 class	8 classes
Formal care hours	-0.51 (0.017)***	-0.42 (0.081)***
Market work hours	-0.17 (0.01)***	-0.08 (0.021)***
Household production hours	0.03 (0.003)***	0.02 (0.008)***

Gross Child Care Price Increased by 10%		
No. of classes	1 class	8 classes
Formal care hours	-0.78 (0.027)***	-0.71 (0.159)***
Market work hours	-0.25 (0.014)***	-0.16 (0.060)***
Household production hours	0.04 (0.005)***	0.03 (0.013)***

Standard errors in the parentheses, significance levels: 90*, 95**, 99***.

Part II

EFFECTIVENESS OF FISCAL STIMULI FOR WORKING PARENTS

4

THE EFFECTIVENESS OF FISCAL STIMULI FOR WORKING PARENTS

The content of this chapter is based on joint work with Henk-Wim de Boer and Egbert Jongen that previously appeared in the *CPB Discussion Paper Series* (286).

4.1 INTRODUCTION

In this chapter we compare the effectiveness of fiscal policies targeted at working families with children which aim to promote parental labour participation. There are large differences in the mix of fiscal support for these families across countries. For example, Scandinavian countries direct much of their public support for working parents to childcare subsidies (OECD, 2014; Kleven, 2014), whereas the US and Canada rely more on in-work benefits to support this group (Immervoll and Pearson, 2009). Although these policies in part differ in their objectives, e.g. promoting skill formation among disadvantaged children versus income support for disadvantaged families, a common goal is that they aim to stimulate employment. There is a large body of literature studying the employment effects of childcare subsidies (and pre-kindergarten and pre-school programs),[1] and there is large body of literature studying the employment effects of in-work benefits for families with children.[2] However, we know very little on the relative effectiveness of these policies in terms of additional employment per additional dollar or euro spent, and hence the policy mix that works best for employment. Furthermore, there are large differences across countries when it comes to the targeting of these policies. For example, in-work benefits for families in the US and the UK are primarily targeted at low incomes (Brewer, Francesconi, Gregg, and Grogger, 2009), whereas in-work benefits for families in the Netherlands are targeted more at middle and high incomes (see below). Targeting childcare subsidies and in-work benefits at working parents with low incomes may cause a loss in efficiency. This, however, depends on the relative importance of labour supply responses on the extensive (participation) and intensive (hours worked per employed) margin

1 See Blau (2003) for an excellent overview, and Lokshin (2004), Tekin (2007), Baker, Gruber, and Milligan (2008), Cascio (2009), Havnes and Mogstad (2011) and Fitzpatrick (2012) for some recent analyses.

2 Two major in-work benefit programs that have received much attention in the literature are the EITC in the US and the WFTC in the UK. See Hotz, Mullin, and Scholz (2010) and Brewer and Browne (2006) and the references therein for the impact of the EITC and WFTC on employment, respectively.

(Saez, 2002). Also here, we know very little on the efficiency loss (if any) of targeting income support more at working parents with low incomes rather than middle and high incomes.

We offer a systematic analysis of the effectiveness of childcare subsidies and in-work benefits for families with children in terms of labour participation. Specifically, we consider how these policies compare to each other in terms of additional public spending per additional (fulltime equivalent) employed, where we show that it is crucial to take into account the effects of behavioural responses on the government budget. Furthermore, we consider to what extent targeting these fiscal policies at different income groups affects their effectiveness, to study the equity-efficiency trade-off for these policies.

To study the effectiveness of fiscal policies targeted at working parents we develop and estimate a structural model of parental labour supply and childcare demand in the Netherlands. We use a large and rich administrative household dataset[3] for the period 2006–2009 to estimate the preferences of couples with a youngest child 0–3 years of age (pre primary school age) and couples with a youngest child 4–11 years of age (primary school age). Specifically, we estimate the preferences using a static discrete choice model for the simultaneous choice of labour supply by the mother and the father, and the use of childcare. [4] An advantage of the discrete choice approach is that it does not require convex or piece-wise linear budget sets, so that we can take all the complexities of the tax-benefit system into account (Van Soest, Das, and Gong, 2002). Furthermore, quasi-concavity of preferences need not be imposed ex ante, and therefore coherency of the model does not implicitly limit the range of behavioural responses that can be obtained (MaCurdy, Green, and Paarsch, 1990). We model unobserved heterogeneity using the latent classes approach as outlined in Train (2008) and Pacifico (2012), and recently applied to a model with maternal labour supply and childcare choices by Apps, Kabátek, Rees, and Van Soest (2012). Latent classes are a flexible way of modelling unobserved heterogeneity, which can prove important for inference of the model (Pacifico, 2012). The identification of the structural parameters benefits from a large reform in childcare subsidies and in-work benefits for working parents in the sample period, which generates large exogenous variation in the budget sets. Hence, we go beyond an identification based solely on cross-sectional variation, which may in part be endogenous, resulting in poor identification of

3 The Labour Market Panel (*Arbeidsmarktpanel* in Dutch) of Statistics Netherlands (2012).

4 Building on the work by Van Soest (1995), discrete choice models have become a popular tool for the structural modelling of labour supply, see e.g. Keane and Moffitt (1998), Blundell, Duncan, McCrae, and Meghir (2000), Gong and Van Soest (2002), Blundell and Shephard (2012) and Bargain, Orsini, and Peichl (2014). For an overview of discrete choice models that explicitly include childcare see Blau (2003). Recent applications include Lokshin (2004), Kornstad and Thoresen (2006) Kornstad and Thoresen (2007), Tekin (2007), Blundell and Shephard (2012), Gong and Breunig (2012) and Apps, Kabátek, Rees, and Van Soest (2012).

the structural parameters and a wide range of potential biases (Blau, 2003). The reform also allows us to do a 'reality check' (Blundell, 2012) on the behavioural responses of the structural model, by comparing the simulated responses to the reform with the findings of a difference-in-differences analysis on the same reform but using a different data set (Bettendorf, Jongen, and Muller, 2012).[5]

Our main findings are as follows. First, with latent classes, the structural model predicts labour supply responses for fiscal reforms over the period 2005–2009 very much in line with the results from the difference-in-differences analysis on the same reforms. When we do not allow for latent classes, the structural model predicts behavioural responses that are smaller than those of the difference-in-differences analysis. Second, we find that a universal increase in childcare subsidies is more effective in raising labour participation than in-work benefits targeted at both primary and secondary earners, and about equally cost-effective as a universal increase of in-work benefits for secondary earners. Third, we find that the effect of childcare subsidies on total hours worked is not much lower when targeted more at low incomes than when targeted at middle and high incomes.[6] However, the so-called knock-on effects, changes in government expenditures and receipts due to behavioural changes, are more favourable when we target childcare subsidies more at middle and higher incomes, making them more cost-effective. Fourth, the most cost-effective fiscal stimulus for working parents is an in-work benefit targeted at secondary earners that rises with income. This provides incentives both on the extensive and intensive margin to a group of workers that is relatively responsive on both margins.

The paper makes several contributions to the existing literature. First of all, we have a large policy reform in our data period. This arguably leads to more credible exogenous variation in budget sets than previous structural analyses of labour supply and childcare that relied on cross-sectional variation only. Second, the policy reform also allows for a quasi-experimental check on the behavioural responses of the structural model, and we contribute to a small but growing literature that evaluates the performance of structural models

5 Our approach satisfies all the requirements set out by Meghir and Phillips (2010, p. 227) "[E]stimating incentive effects in a convincing way thus requires us to find solutions to all these problems at the same time. This calls for a sufficiently flexible approach, that allows for fixed costs of work, does not impose theory a priori everywhere in the sample (thus in a sense increasing model flexibility), uses exogenous changes to work incentives to identify their effect, and allows for taxes and benefits. This is of course a large set of requirements, but all have been shown to be important empirically; in our review of empirical results we will use these criteria to judge the value of the estimates."

6 The case for targeting childcare subsidies at low incomes is reinforced when participation in childcare benefits children from low incomes more than children from middle and high incomes, as suggested by e.g. Blau and Currie (2006) and Havnes and Mogstad (2014).

by comparing simulated policy responses with the results from quasi-experimental studies (Todd and Wolpin, 2006; Hansen and Liu, 2011; Geyer, Haan, and Wrohlich, 2014). Third, with the structural model, we can also study a number of issues that were not possible in the quasi-experimental analysis. We decompose the labour participation effect of the 2005–2009 reform package into the effect of changes in childcare subsidies and the effect of changes in in-work benefits. We also do counterfactual policy analysis, including a prediction for the labour participation effects of a recent cut in childcare subsidies in the Netherlands. Because our structural model is fully integrated with a detailed tax-benefit calculator, we can study the effectiveness of fiscal stimuli for working parents in terms of additional employment generated per additional public dollar or euro spent. The integrated model also allows us to go beyond back-of-the-envelope calculations on the effectiveness of different types of family policies using population averages for e.g. taxes and childcare subsidies as in Blau (2003) and Lokshin (2004). Although we focus on the impact of policy reforms in the Netherlands, we argue that our findings are also relevant for the effectiveness of these policies in other developed OECD countries. Indeed, the participation rate of mothers and fathers in the Netherlands, as well as public spending on formal childcare and pre-primary education, takes an intermediate position between Scandinavia and Anglo-Saxon countries. Finally, our data set is exceptionally large and rich. Hence, we can identify preferences for a large number of subgroups, including couples with a youngest child that is in primary school. To the best of our knowledge, we are the first to estimate a structural model for labour supply and out-of-school care, next to a model for labour supply and daycare.

The paper is organized as follows. Section 2 describes the policy environment and the evolution of labour market participation by men and women in the Netherlands in an international context. Section 3 develops the structural model and outlines the empirical strategy. Section 4 describes the data. Section 5 presents the estimation results and the corresponding labour supply and childcare elasticities. In this section we also present a comparison of the simulated employment effects of the structural model for the 2005–2009 reform package with the estimated employment effects of a quasi-experimental study. In Section 6 we use the structural model to compare the effectiveness of different fiscal stimuli for working parents. In Section 7 we simulate the employment effects of recent cuts in childcare subsidies. Section 8 concludes. Supplementary material is given in the appendix.

4.2 LABOUR MARKET AND POLICY ENVIRONMENT

In the mid 1970s, the participation rate of women (15–64 years of age) in the Netherlands was rather low by international standards, close

to 30% (OECD, 2013).[7] However, following the economic crisis in the early 1980s, the participation rate of women in the Netherlands started to rise. The rise in participation by mothers of young children was particularly strong (Euwals, Knoef, and van Vuuren, 2011). By 2004, the Netherlands, with a participation rate of women close to 70%, took an intermediate position between the higher participation rates in e.g. Norway and Sweden, and the lower participation rates in e.g. the US and the UK.[8]

The participation rate of men in the Netherlands dropped from the mid 1970s to the mid 1980s. In the face of adverse labour market conditions, many men were sent into early retirement and disability. However, in the 1990s and 2000s the generosity of early retirement and disability was cut back, and participation rates returned to levels comparable to other developed OECD countries.[9]

To further promote the labour participation (in persons but also in hours worked per week) by families with children, and of mothers in particular, the Dutch government implemented a series of reforms over the period 2005–2009. Following a brief introduction into the pre-reform childcare market in the Netherlands, below we give a short historical account of the policy changes over the period 2005–2009.

Children in the Netherlands go to primary school when they turn 4 (the legal obligation is in place from the age of 5 on), and most children are 12 years old when they go to secondary school. Before the age of 4, children can go to centre-based daycare, so-called playgroups (*peuterspeelzalen*) and informal care. Before the introduction of the Law on Childcare (*Wet kinderopvang*) in 2005, centre-based daycare was subsidized at varying rates.[10] The majority (76%) of places was subsidized directly by employers and local governments.[11] These places had lower effective parental fees than so-called 'unsubsidized' places (24%), the costs of which were however partly tax deductible for parents. To qualify for the subsidies and tax deduction, both parents for two-parent households and one parent for single-parent households need to work. The total enrollment rate of children 0–3 years of age in centre-based care was 25% in 2004 (see Figure 4.1). Next to centre-based care, a large number of children also go to playgroups.

7 This section draws heavily on Bettendorf, Jongen, and Muller (2012).

8 Whereas the participation rate of women in the Netherlands has converged to other well-developed OECD countries, there remains a sizeable and stable gap in hours worked by employed women (OECD, 2013). In 2004, employed women in the Netherlands worked on average approximately 24 hours per week, while their counterparts in other OECD countries worked 5 to 10 hours per week more. Indeed, in 2004, the share of women working part-time in the Netherlands was 60%, by far the largest share in the OECD (OECD, 2013).

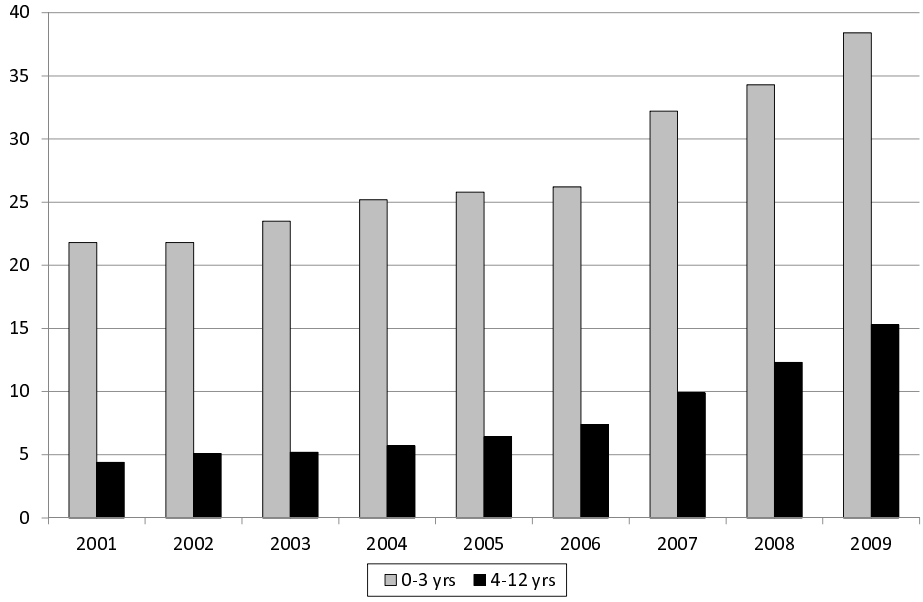
9 Hours worked per week by employed men (36 hours per week in 2004) is also somewhat lower in the Netherlands than in other well-developed OECD countries, but the difference is much less pronounced than for women (OECD, 2013).

10 All the data on the use of formal childcare in this section are from Statistics Netherlands (http://statline.cbs.nl).

11 The subsidy is per hour of formal childcare.

This is part-time care for less than 4 hours per day, mostly used by families in which one of the parents does not work. Playgroups are not a subsitute for centre-based care as they do not cover enough hours of care for the parents to work. The enrollment rate of children 0–3 in playgroups was also close to 25%. Children that are in primary school (4–12 years of age) can go to centre-based out-of-school care and informal care. Similar to daycare, before the introduction of the Law on Childcare, subsidized and unsubsidized centre-based out-of-school care places co-existed, where the costs of unsubsidized places were partly tax deductible for parents. The pre-reform enrollment rate of 4–12 year olds in centre-based care was 6% in 2004.

Figure 4.1: Share of children in formal childcare (in %)



Source: Statistics Netherlands.

The series of reforms started with the introduction of the Law on Childcare in 2005. This law unified the subsidies for childcare places. From 2005 onwards, all formal places qualified for the same subsidy from the central government. This increased the subsidy somewhat for parents with children going to an unsubsidized place before 2005. Care by childminders, at the home of the childminder or of the children, also became eligible for subsidies under this law. But the unification of the subsidies and the extension to care by childminders had only a minor effect on public spending on formal childcare. Indeed, the subsidy was actually reduced somewhat for the highest incomes[12], public spending actually fell slightly from 2004 to 2005, see Table 4.1.

More important were the changes that followed in 2006 and 2007. In these years the subsidy rate was increased drastically, in particular in 2007. Figure 5.1 shows the changes in the parental contribution rate

12 See Plantenga, Wever, Rijkers, and de Haan (2005).

Table 4.1: Public spending on childcare and in-work benefits for parents (millions of euro)

Year	2002	2003	2004	2005	2006	2007	2008	2009
Childcare subsidies	725	755	1,028	1,001	1,343	2,058	2,825	3,034
In-work benefits for parents	410	460	738	830	871	984	971	1,290
– *Combinatiekorting*[a]	410	460	479	484	314	324	247	0
– *Inkomensafhankelijke Combinatiekorting*[b]	0	0	259	346	557	660	724	1,290

Source: Ministry of Finance (2010) and own calculations (imputation of employers' contribution for childcare up to 2007 with data from the Ministry of Social Affairs and Employment (personal communication) and split of the in-work benefits for parents in its two components using the MIMOSI model of CPB). [a]The *Combinatiekorting* applies to primary earners, secondary earners and working single parents with a youngest child up to 12 years of age. [b]The *Inkomensafhankelijke Combinatiekorting* applies to secondary earners and working single parents with a youngest child up to 12 years of age.

for the 'first child'.[13] The parental fee depends on the income of the household. In all years, households with the lowest income receive the highest subsidy (up to 96% of the full price). For the lowest income households the subsidy rate hardly changed. For the middle income households the subsidy rate went up by 20 to 40%-points, whereas the increase in the subsidy for the highest income households was somewhat smaller than for middle income households. On average, the parental cost share in the full price dropped from 37% in 2005 to 18% in 2007.[14,15] Next to the drop in parental fees, from 2007 onwards schools were obliged to act as an intermediary for parents and childcare institutions to arrange out-of-school care.

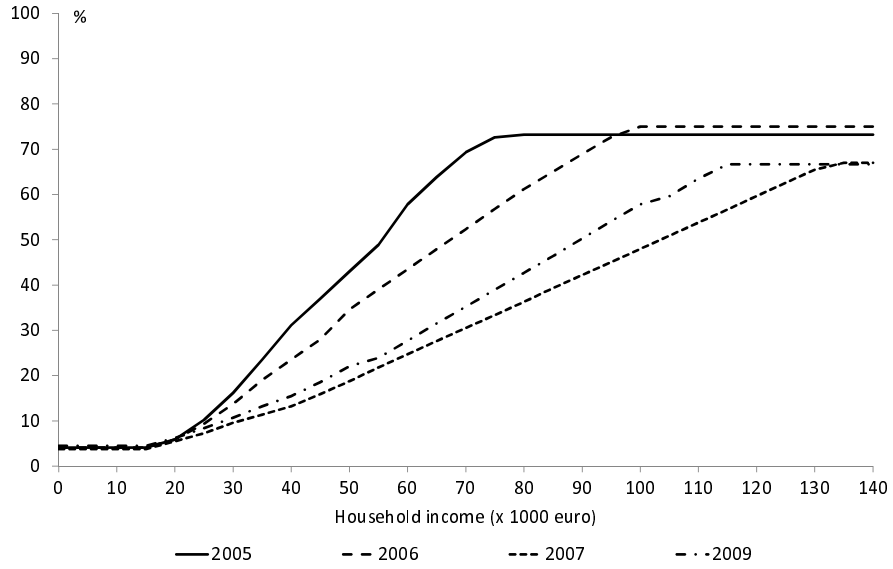
In 2008 there were virtually no changes in childcare subsidies. 2009 then witnessed a partial reversal of the increase in childcare subsidies, as subsidies were cut back somewhat, see again Figure 5.1. Over the period 2005–2009, public spending on formal childcare went from 1 to 3 billion euro. By 2009, with public spending on childcare and pre-primary education of 0.5% of GDP, the Netherlands took an intermediate position between Sweden and Norway that spent respectively 1.4 and 1.2% of GDP on these policies on the one hand, and the US and Canada that spent just 0.4 and 0.2% of GDP on these policies on the other (OECD, 2014). Figure 4.1 shows the corresponding rise in the use of formal childcare over the period 2001–2009 in the Netherlands. Following the steep drop in the parental fee in 2006 and 2007, there was a steep rise in the use of formal childcare, both for children 0–3 years of age (daycare) and for children 4–12 years of age (out-of-school care).

13 The Tax Office defines the first child as the child for which the parents have the highest childcare expenditures. For most households the first child is the youngest child since more hours are needed for daycare (0–3 years of age) than for out-of-school care (4–11 years of age).

14 Source: Tax Office data provided by the Ministry of Social Affairs and Employment (personal communication).

15 Despite the steep increase in the subsidy rate, the average prices of formal childcare places grew more or less in line with the CPI.

Figure 4.2: Parental contribution rate for the first child



Source: own calculations using publicly available subsidy tables.

The period 2005–2009 also witnessed a number of changes in in-work benefits for working parents. Figure 5.3 shows the level of the Combination Credit (*Combinatiekorting*) per year over this period. All working parents with a youngest child less than 12 years of age qualified for the Combination Credit.[16] Furthermore, the in-work benefit was independent of earned income, provided earned income was above a certain (low) threshold. This benefit was introduced in the major tax reform of 2001, but was phased out over the period 2005–2009. There was a reduction in 2006, and then a smaller reduction in 2008 before it was eventually abolished in 2009.

Figure 5.4 shows the level of the Income-Dependent Combination Credit (*Inkomensafhankelijke Combinatiekorting*) per year by earned income over the period 2005–2009.[17] Secondary earners (and single parents) qualify for this in-work benefit, but the primary earners of secondary earners do not. This benefit was introduced in 2004. Up to 2008, there was a gradual increase in the tax credit, and the credit did not depend on earned income (again provided that earned income exceeded a certain threshold). In 2009 this tax credit became income dependent, with a phase-in rate of 3.8% for income above the threshold. The maximum credit in 2009 was 1,765 euro, where the maximum was reached at a gross individual income of 30,803 euro.[18]

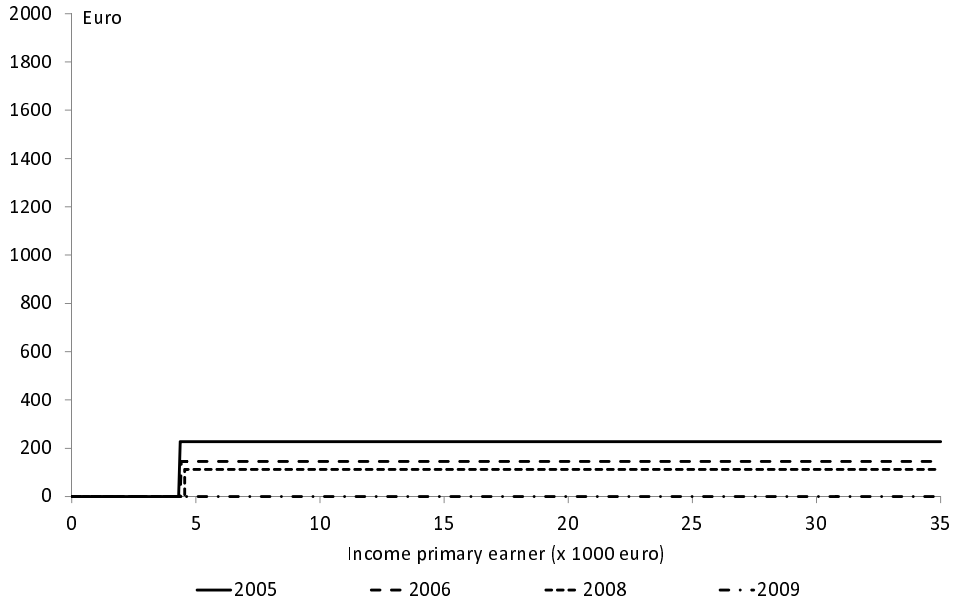
As Figure 5.3 and 5.4 show, there was a shift from the Combination Credit, for which both primary and secondary earners were eligible, to the Income-Dependent Combination Credit, for which only

16 The name refers to the combination of work and care.

17 Up to 2008 the *Inkomensafhankelijke Combinatiekorting* was called the *Aanvullende Combinatiekorting* (Additional Combination Credit).

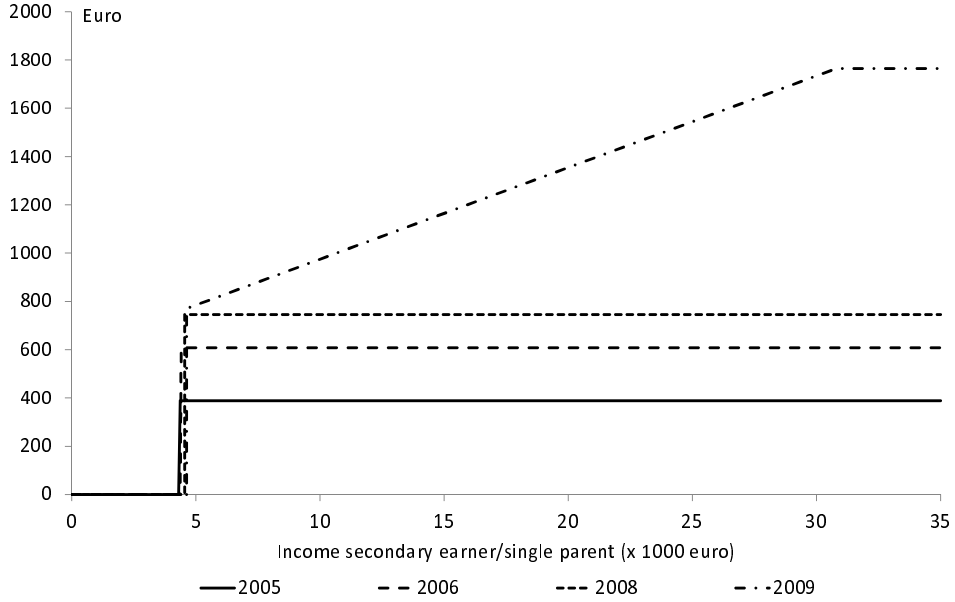
18 For comparison, in 2009 the minimum wage of a fulltime worker was 16,776 euro.

Figure 4.3: Annual in-work benefit for primary and secondary earners with children



Source: Tax Office.

Figure 4.4: Annual in-work benefit for secondary earners with children



Source: Tax Office.

secondary earners[19] (typically mothers) were eligible. Indeed, public expenditures on the Combination Credit dropped from 484 million euro in 2005 to 0 in 2009, whereas public expenditures on the Income-Dependent Combination Credit rose from 346 million euro in 2005 to 724 million euro in 2008, and then to 1,290 million euro in 2009 as the income dependent part was added, see Table 4.1.[20] The motivation for these changes in in-work benefits was that secondary earners were believed to be more responsive to financial incentives than primary earners, and that policymakers wanted to stimulate mothers in the Netherlands to work more hours per week.

The reforms we simulate below are motivated by the policy reforms we witnessed in the Netherlands over the past decade. Indeed, in the analysis below we want to determine what policy seems to be the most effective in terms of labour participation. Furthermore, we want to quantify the trade-off between efficiency and equity by considering the redistributional and efficiency effects of targeting childcare subsidies and tax credits at different income groups.

4.3 STRUCTURAL MODEL AND EMPIRICAL METHODOLOGY

Households are assumed to maximize a unitary household utility function. The systematic part of household utility, U^s, depends on disposable income y, hours worked by the male h_m, hours worked by the female h_f, and hours of formal childcare used c.[21] The functional form of U^s is log-quadratic,

$$\begin{aligned} U^s(\nu) &= \nu'\mathbf{A}\nu + \mathbf{b}'\nu + \mathbf{d}'\mathbf{1}[\mu > 0], \\ &\quad \nu = (\log(y), \log(1 - h_m/T), \log(1 - h_f/T), \log(c)), \\ &\quad \mu = (h_m, h_f, c), \end{aligned} \tag{4.1}$$

with $\mathbf{A}$ being a symmetric matrix of quadratic coefficients and $\mathbf{b}$ being a vector of linear coefficients corresponding to the vector of the aforementioned variables ν.[22] The vector $\mathbf{d}$ captures fixed costs of work for men and women and fixed costs of using formal childcare. Since these fixed costs are specified in the utility metric, they represent an amalgamation of different factors such as intrinsic disutility from

19 And single parents.

20 Unfortunately, we could not find internationally comparable data on total public spending on in-work benefits for families with children.

21 Unfortunately we do not observe informal childcare in our administrative dataset. In a robustness check we include a proxy for the use of informal childcare as an additional argument in the utility function. In the robustness check we assume that the total demand for childcare equals $c_{tot} = max((h_m + h_f - T), 0)$. We then use the following proxy for informal childcare $c_{inf} = max((c_{tot} - c), 0)$. Using this extended specification leads to similar labour supply and childcare elasticities, see the appendix.

22 Note that the parental work variables h_m and h_f in the vector ν have been transformed into indicators of leisure utilization, representing the fraction of weekly time endowment T which is spent on activities unrelated to work (including self-provided childcare and household maintenance).

work, or market frictions and other costs related to job search and childcare use. We allow for preference variation through observed individual and household characteristics $\mathbf{x}_2$, $\mathbf{x}_3$ and $\mathbf{x}_4$ in parameters b_2, b_3 and b_4

$$b_2 = \mathbf{x}'_2\mathbf{fi}_2, \quad b_3 = \mathbf{x}'_3\mathbf{fi}_3, \quad b_4 = \mathbf{x}'_4\mathbf{fi}_4, \tag{4.2}$$

which are the linear utility terms in leisure of the male, leisure of the female, and hours of formal childcare, respectively. The same variation is also allowed for the fixed costs parameters $\mathbf{d}$ (for a full list of covariates used, see appendix 4.D).

The budget constraint takes the following form

$$y = w_m h_m + w_f h_f - T(w_m, h_m, w_f, h_f; q) - TC(p_c, c; q) + S(p_c, c, y_t; q), \tag{4.3}$$

where w_m and w_f denote the gross hourly wage for the male and the female,[23] $T(.)$ denotes taxes and employees' premiums, q denotes individual and household characteristics, $TC(.)$ is the total cost of formal childcare, with p_c denoting the price per hour of formal childcare, and $S(.)$ is the childcare subsidy, which depends on the hourly price of formal childcare, the hours of formal childcare, taxable income y_t and household characteristics like the age distribution of the children.

Our econometric specification is based on a discrete choice model. Parents choose their preferred combination of hours of work and the hours of formal childcare from a finite set of alternatives $j \in \{1, ..., J\}$. Disposable household income depends on these choices, rising in hours worked and falling in formal childcare demanded. For workers we observe gross wages which are used to compute the work-related part of income for each alternative in their choice sets. For non-workers we estimate a Heckman-type wage equation which is used to simulate their wages. We account for wage heterogeneity by taking multiple draws from the wage error distribution. Similarly, for households that use formal childcare we use observed hourly prices of formal childcare, and for non-users we simulate hourly prices using the same estimation strategy as for hourly wages. A detailed description of both simulation exercises can be found in the appendix.

Next to the systematic part $U_s(v_j)$, the utility function also contains alternative-specific stochastic terms ε_j:

$$U(v_j) = U^s(v_j) + \varepsilon_j. \tag{4.4}$$

The stochastic terms are assumed to be i.i.d. across alternatives, and to be drawn from the Type 1 Extreme Value distribution. This leads to a multinomial logit specification of the discrete choice model.

We also allow for the possibility that families which are observationally equivalent might have different tastes for work and formal childcare. We assume that there is a finite number K of latent household

23 We assume that the gross hourly wage does not depend on the hours worked.

classes (or types), with households having homogeneous preferences within each class but heterogeneous preferences across classes. In practice, this means that we estimate a finite mixture model with K parametrizations of the utility function, corresponding to K distinct subsets of our data. All the preference parameters therefore become class-specific, which is equivalent to the assumption that they are drawn from a mass-point distribution.[24] The full set of parameters to be estimated is then

$$\theta = (\theta_1, ..., \theta_K) = (\mathbf{A}_1, \mathbf{b}_1, \mathbf{d}_1, ..., \mathbf{A}_K, \mathbf{b}_K, \mathbf{d}_K). \tag{4.5}$$

Since the classes are by definition unobservable, we cannot determine whether a given household belongs to a specific class or not. Instead, we have to construct household-level probabilities of class membership $P_i(class = k)$, which reflect how likely is household i to be driven by the preferences corresponding to class k, conditional on the household's choices and other observable characteristics. These probabilities are then used as individual weights for a set of class-specific multinomial logit models with separate parameter vectors θ_k. The resulting log-likelihood function of the finite mixture model has the following form

$$L = \sum_{i=1}^{I} \log \left(\frac{1}{R} \sum_{r=1}^{R} \sum_{k=1}^{K} P_i(class = k) \cdot \sum_{j=1}^{J} \left(\frac{\exp\left(U_{ij}^{s}(\nu_r, \theta_k)\right)}{\sum_{j'=1}^{J} \exp\left(U_{ij'}^{s}(\nu_r, \theta_k)\right)} \cdot D_{ij} \right) \right), \tag{4.6}$$

where R denotes the number of draws from the estimated wage and price equation for non-workers and non-users of formal childcare.[25] D_{ij} is an indicator function which takes the value 1 for the observed choice, and zero otherwise.

To solve the model, we use the EM algorithm, as proposed by Train (2008). This approach has been chosen since the likelihood frontier is likely to violate global concavity, which renders the solution by conventional methods based on maximum likelihood practically infeasible.

24 Limiting the distributional assumptions on unobserved heterogeneity by using mass points was pioneered by Heckman and Singer (1984). Recently, Train (2008) introduced a tractable way of estimating latent class discrete choice models using the EM algorithm. For a discussion of the benefits of latent class models within the domain of structural labour supply modelling, see Apps, Kabátek, Rees, and Van Soest (2012). For an overview of their implementation and potential computational improvements, see Kabátek (2013).

25 The number of draws in our specification is 10, and it is kept relatively low to limit the computational complexity of the model. We argue that this is sufficient since the unobserved component in the childcare price equation is negligible compared to the actual values - reflecting the fact that the majority of childcare centres charge the same price for their services. Increasing the number of draws does not qualitatively change predictions of our model.

4.4 DATA

We use the Labour Market Panel (in Dutch: *Arbeidsmarktpanel*) of Statistics Netherlands (2012). The backbone of the Labour Market Panel are the annual observations of the Labour Force Survey (in Dutch: *Enquete Beroepsbevolking*) for the period 1999–2009, which contains the education level of adult members of the household. Statistics Netherlands supplements this data set with three additional data sources. First, administrative data from municipalities for the period 1999–2009 (in Dutch: *Gemeentelijke Basisadministratie*) that contains information on individual and household characteristics like age, ethnicity, ages of the children and area of residence. Second, administrative data from the Social Statistical Panel for the period 1999–2009 (in Dutch: *Sociaal Statistisch Bestand*) on hours worked and gross income. The hours worked are observed both in their nominal value and also expressed as a share of the full-time work equivalent spent in a given job. Third, administrative data on formal childcare from the Formal Childcare Database of the Tax Office for the shorter period 2006–2009 (in Dutch: *Wet Kinderopvangtoeslag*). With respect to formal childcare, a distinction is made between daycare (children 0–3 years of age[26]) and out-of-school care (children 4–11 years of age).

From the Labour Market Panel we make the following selections to arrive at the sample we use in the estimations. Childcare subsidies are available to parents up to the point where the child goes to secondary school. Most children are 12 when they go to secondary school[27], and therefore we restrict the sample to households with a youngest child 0 up to and including 11 years of age. Because we only have data on the use of formal childcare for the period 2006–2009, we further need to restrict the sample to this period.

We only model the labour supply choice of couples, and hence also drop single parents. Next, we exclude couples in which at least one parent is either self-employed or has multiple sources of income, because we can not determine their budget constraint.[28] Furthermore, we exclude couples in which at least one of the partners is on disability or unemployment benefits, assuming that they are constrained in their labour supply choice. After these selections are made, we further drop households with missing information on individual or household

26 Maternity leave in the Netherlands is rather short, 3 months after the birth of the child, which can be supplemented with 3 months of parental leave for which the replacement rate is rather low however (OECD, 2014). Hence, we also include parents with a youngest child less than 1 years old in the analysis.

27 We do not observe whether a child is in secondary school or not.

28 The self-employed account for 8.1% of working mothers, which is a lower share than the one observed in the Australian data (13%). We argue that unlike in the Australian case, here we are not fundamentally limiting the flexibility of maternal work choices by excluding self-employed mothers. The Dutch labor markets are known to be very accomodating to part-time work appointments, so that mothers do not need to opt for self-employment in order to reach out for more flexible labor engagement.

characteristics. This leaves us with approximately 60 thousand observations (households times periods in the sample). Given the large set of discrete choices we allow (see below), and the large set of preference parameters for each latent class, estimating the preference parameters results in a considerable computational burden. We therefore take a random subsample of 15%.[29] This leaves us with 4,170 observations for couples with a youngest child 0–3 years of age, and 5,013 observations for couples with a youngest child 4–11 years of age.

Table 4.2 gives descriptive statistics of our sample. Fathers in our sample are on average a few years older than mothers. Fathers and mothers in our sample are predominantly born in the Netherlands, and most of them have a level of education classified as middle. Furthermore, whereas fathers with a youngest child 0–3 years of age are slightly more likely to be higher educated than fathers with a youngest child 4–11 years of age, mothers with a youngest child 0–3 years of age are considerably more likely to be higher educated than mothers with a youngest child 4–11 years of age (a cohort effect). The majority of households lives in smaller cities and towns (<150,000 inhabitants). There is a considerable gap in the gross hourly wage between fathers and mothers, with fathers earning on average 4 to 6 euros per hour more than mothers in couples with a youngest child 0–3 and 4–11 years of age, respectively. The labour participation rate is much higher for fathers than for mothers. Furthermore, the participation rate of mothers with a youngest child 0–3 is higher than the participation rate of mothers with a youngest child 4–11. Finally, households with a youngest child 0–3 years of age are more likely to use formal childcare than households with older children. 50% of the households with a youngest child 0–3 years of age sends their children to formal childcare, compared to just 13% for households with a youngest child 4–11 years of age. A typical school day is from 8:30 to 15:00, and many families are able to cover the remaining hours with parental time or informal care. This is also reflected in the average hours of formal childcare used per week by households that do use formal childcare. It remains to be added that the differences between the statistics observed within the two samples of mothers are also likely to reflect the cohort effects. The education levels of women with older children are considerably lower than those attained by women with children aged 0–3. Part of the differences observed might be hence related to the educational differences, and potentially also the different labor prospects stemming therefrom. These cohort effects sohuld be also borne in mind when interpreting the estimation results for the both samples.

29 We have tested the stability of the preferences and the elasticities using different subsample sizes. Moving from smaller to larger sample sizes, preferences and elasticities appear to stabilize once we take a 15% subsample.

Table 4.2: Descriptive statistics by sex and age of the youngest child

	Men				Women			
	0-3 yrs		4-11 yrs		0-3 yrs		4-11 yrs	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	36.8	4.90	43.3	5.10	34.1	4.40	40.8	4.60
Native	0.84	0.36	0.86	0.35	0.84	0.37	0.84	0.37
Western immigrant	0.08	0.27	0.07	0.26	0.09	0.28	0.09	0.28
Non-Western immigrant	0.08	0.26	0.07	0.25	0.07	0.27	0.08	0.26
Lower educated[a]	0.19	0.39	0.21	0.41	0.14	0.34	0.22	0.42
Middle educated[a]	0.44	0.50	0.44	0.50	0.46	0.50	0.51	0.50
Higher educated[a]	0.38	0.48	0.35	0.48	0.40	0.49	0.26	0.44
Large city[b]	0.16	0.37	0.16	0.36	0.16	0.37	0.16	0.36
Small city[b]	0.84	0.37	0.84	0.36	0.84	0.37	0.84	0.36
Hourly gross wage	20.2	10.0	22.2	11.2	16.3	6.30	16.1	7.60
Participation rate	0.96	0.19	0.95	0.21	0.82	0.39	0.75	0.43
Hours worked per week[c]	38.7	5.20	38.7	5.50	23.0	8.20	21.2	8.50
Using formal childcare[d]	0.50	0.50	0.13	0.34	0.50	0.50	0.13	0.34
Hours formal childcare per week[e]	27.1	16.2	14.4	11.1	27.1	16.2	14.4	11.1
Number of observations	4,170		5,013		4,170		5,013	

[a]Education is classified as follows (using the Dutch abbreviations): i) lower educated = BO and VMBO, ii) middle educated = MBO, HAVO and VWO, iii) higher education = HBO and WO. [b]A city is defined as large (small) when it has 150,000 inhabitants or more (less than 150,000 inhabitants). [c]Hours worked per week per employed. [d]The share of households using formal childcare is higher than the share of children in formal childcare in Figure 4.1. First, Figure 4.1 includes more households who are not eligible for childcare subsidy such as households with unemployment or disability benefits. Second, households with many children use less formal childcare on average. [e]Hours of formal childcare per week per household using formal childcare.

For our discrete choice model we discretize the data. Men and women are both allowed to choose from 6 labour supply options. Labour supply is discretized in 0 to 5 days, where each day equals 8 hours.[30] For childcare, we allow for 0, 1, 2 and 3 days, where the data show that a typical day in a daycare centre equals 10 hours,[31] and a typical day in out-of-school care equals 5 hours.[32] The full choice set for each household is 6 X 6 X 4 = 144 alternatives.

To determine disposable household income in each discrete option for labour supply and formal childcare we use the MIMOSI model (Romijn, Goes, Dekker, Gielen, and van Es, 2008). MIMOSI is the official tax-benefit calculator of the Dutch government for the (non-behavioural) analysis of the redistributional and budgetary effects of reform proposals. MIMOSI allows for a very accurate calculation of the budget constraints. Indeed, it takes into account all (national[33]) taxes, social security premiums, and income independent subsidies and tax credits. Furthermore, MIMOSI also calculates the childcare subsidy applicable for each household in each option. The subsidy depends on the full hourly price of childcare per type of childcare (e.g. daycare or out-of-school care) up to a maximum price beyond which parents receive no additional subsidy, household income (subsidies are lower for higher incomes), the number of children (the subsidy is higher for the second, third etc. child in formal childcare), and whether or not both parents work (both parents need to work to receive the subsidy[34]). Income that enters the household utility function is disposable household income defined as gross household income plus childcare subsidies minus taxes, employees' premiums (for the employed), the nominal health care fee, and expenditures on formal childcare.[35] We ensure that household disposable income (excluding childcare costs and childcare subsidies) can not fall below the social assistance level for couples with children. For each discrete option we also calculate the net transfer from the household to the government (positive or negative). This allows for an accurate calculation of the net budgetary costs of the reforms excluding and including behavioural responses. This is crucial for the comparison of the effectiveness of different fiscal stimuli for working parents.

30 Classified as: $0 \in [0,5), 8 \in [5,13), 16 \in [13,21), 24 \in [21,29), 32 \in [29,37), 40 \in [37,\infty)$.

31 Classified as: $0 \in [0,0], 10 \in [0,15), 20 \in [15,25), 30 \in [25,\infty)$.

32 Classified as: $0 \in [0,0], 5 \in [0,7.5), 10 \in [7.5,12.5), 15 \in [12.5,\infty)$.

33 In the Netherlands local taxes account for only a small portion of total taxes (3.3% in 2007, European Union (2014)).

34 When one of the partners becomes unemployed, they are still eligible for childcare subsidies for a limited period of time.

35 Disposable income in the estimations and simulations is in 2006 prices. We use the CPI to convert prices in later years to 2006.

4.5 ESTIMATION RESULTS

We estimate the preferences separately for couples with a youngest child 0–3 years of age, and for couples with a youngest child 4–11 years of age. This is to acknowledge that there can be non-trivial differences in childcare requirements and labour supply incentives faced by the two groups of households (Bernal, 2008).

As discussed in Section 5.2, to account for unobserved heterogeneity, we allow each subpopulation to consist of a number of latent classes. In order to assess how many latent classes should be used, we have estimated a set of models allowing for 1, 2, 3 or 4 latent classes (the model with one class being a homogenous specification). The key variables of interest, the labour supply and formal childcare elasticities, prove to be fairly stable for specifications with two and more latent classes. These can be found in the appendix (see Table 4.A.3).

In contrast to Chapter 3, here we do not use Bayesian Information Criteria (BIC) to select the optimal number of latent classes. The BIC criteria which we derived in an earlier version of this paper would lead us to employ highly-stratified models, even though we found that inclusion of all the additional classes had little to no impact on measured elasticities and reform simulations.[36]. Therefore we have decided to use the most parsimonious model which would generate elasticities qualitatively similar to those of more-stratified models, opting for the specification with 2 latent classes. Another way to select the optimal number of latent classes would be to use a problem-specific version of the information criterion, following the work of Claeskens and Hjort (2003). The authors advocate the use of Focused Information Criterion, which aims to assess the stability of a problem-specific parameter of interest (such as wage elasticity), rather than the stability of penalized log-likelihood as is the case with BIC.

The estimated preference parameters and aggregate class shares for the models with 2 latent classes can be found in the appendix (Table 4.A.5 and 4.A.6). However, rather than interpreting the individual coefficients, we focus on elasticities derived from the estimated structural parameters. First, consider the labour supply elasticities in Table 4.3. For an increase in the gross hourly wage of the men, we find a total hours worked elasticity for men ('Labour supply men') of 0.06 (youngest child 4–11) and 0.08 (youngest child 0–3), where most of the response is on the decision whether or not to participate ('Extensive margin') and not on the decision on how many hours per week to

36 This undesirable outcome follows from the mechanics of BIC: The second term in the BIC equation 3.11 (which penalizes log-likelihood for the loss of degrees of freedom) becomes less important as the number of observations n grows larger. Accordingly, BIC is going to favor extremely stratified models as long as the sample is big enough. However, whether the increased precision of parameter estimates is worth the computational burden of estimating a model with 20 or more latent classes, is an open question with no universal answer.

work ('Intensive margin'). We find a sizeable negative cross-elasticity for total hours worked by women. However, note that women work fewer hours in the base than men. In the end, the overall effect on total hours worked by households is close to zero (not reported). We also find a modest elasticity of the use of formal childcare with respect to the gross hourly wage of the men.

Table 4.3: Gross wage elasticities

	Hourly wage men +1%				Hourly wage women +1%			
	0–3 yrs		4–11 yrs		0–3 yrs		4–11 yrs	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Labour supply men	0.08	0.01	0.06	0.02	–0.05	0.01	–0.04	0.02
– Extensive margin	0.07	0.02	0.04	0.03	–0.02	0.01	–0.01	0.02
– Intensive margin	0.01	0.02	0.02	0.02	–0.03	0.01	–0.03	0.02
Labour supply women	–0.15	0.03	–0.08	0.02	0.40	0.03	0.47	0.03
– Extensive margin	–0.10	0.02	–0.04	0.02	0.25	0.03	0.31	0.04
– Intensive margin	0.00	0.02	0.00	0.01	0.15	0.02	0.16	0.02
Formal childcare	0.11	0.05	0.15	0.07	0.41	0.02	0.77	0.11

Bootstrapped standard errors based on 200 draws.

Turning to the results for an increase in the gross hourly wage of the women, we find much larger own-wage elasticities for women than for men. Indeed, the own-wage elasticity for mothers with a youngest child 0–3 and 4–11 is 0.40 and 0.47, respectively. About two-thirds of the response is on the extensive margin, and about one-third is on the intensive margin.[37] We also find negative cross-elasticities for men, but these cross-elasticities are considerably smaller than for women. Following the larger labour supply response to female wages than male wages, we also find a larger elaticity of the use of formal childcare with respect to the gross hourly wage of women instead of men.

Table 4.4 presents the formal childcare price elasticities. In the first three rows, we consider the elasticity of the use of formal childcare, labour supply by men and labour supply by women with respect to the change in the gross price of formal childcare. We see a substantial negative price elasticity of formal childcare: –0.66 for couples with a youngest child 0–3 years of age and –0.77 for couples with a youngest child 4–11 years of age. There is hardly any effect on the labour

37 Bargain, Orsini, and Peichl (2014) also find that intensive margin responses for women in couples are relatively high in the Netherlands. Indeed, women in the Netherlands are arguably more free to choose their working hours, given the large share of part-time working women in the Netherlands.

Table 4.4: Gross and net price of formal childcare elasticities

	Price of formal childcare +1%			
	0–3 yrs		4–11 yrs	
	Mean	SE	Mean	SE
Gross price elasticities				
Formal childcare	–0.66	0.03	–0.77	0.10
Labour supply men	0.00	0.00	0.00	0.01
Labour supply women	–0.14	0.01	–0.04	0.01
Net price elasticities				
Formal childcare	–0.41	0.02	–0.54	0.07
Labour supply men	0.00	0.00	0.00	0.01
Labour supply women	–0.09	0.01	–0.03	0.01

Bootstrapped standard errors based on 200 draws. The gross price of formal childcare elasticities relate the percentage change in the use of formal childcare and labour supply by men and women to the percentage change in the full price of formal childcare. The net price of formal childcare elasticities relate the percentage change in the use of formal childcare and labour supply by men and women to the percentage change in the parental fee for formal childcare.

supply of men, but a significant negative effect on the labour supply of women. This is particularly true for women with a youngest child 0–3 years of age, who use much more formal childcare than women with a youngest child 4–11 years of age. The next three rows give the same elasticities with respect to the net price of formal childcare or the parental fee of formal childcare. A 1% increase in the gross price leads to more than a 1% increase in the average parental fee in part because a fraction of the parents pays a gross price that is higher than the maximum price for which they can get a subsidy. Hence, these parents have to bear the full 1% rise in the gross price. The net price elasticities are more directly comparable to other studies, that typically focus on the elasticity with respect to the parental fee. These elasticities are somewhat smaller, but still substantial with –0.41 for couples with a youngest child 0–3 and –0.54 for couples with a youngest child 4–11.[38] Our results for the net price elasticity of labour supply by women is in line with the review presented in Blau (2003, p. 492). For the studies that explicitly allow for a formal childcare choice next to a labour supply choice, and hence do not impose a 1-to-1 link between the two, the elasticity of labour supply of women with respect to the net price of formal childcare is relatively low, ranging from –0.09 to –0.20. For mothers with a youngest child 0–3 years of age, we find a similar low elasticity of –0.09. For mothers with a youngest child 4–11

38 For example, in a recent study for Australia, Gong and Breunig (2012, Table 4) calculate a net price elasticity of childcare of –0.22.

Table 4.5: Comparison with DD analysis: policy reforms 2005–2009

	Structural model				DD analysis[b]	
	Childcare	Comb. Credit	Income Dep. CC	Total	Coeff.	SE
Model with latent classes[a]			Changes in levels			
Youngest child 0-3 yrs						
Participation rate women	0.017	-0.005	0.018	0.030	0.020	0.007
Hours worked per week women	0.693	-0.098	0.566	1.185	1.222	0.223
Participation rate men	0.003	-0.002	0.003	0.004	0.006	0.004
Hours worked per week men	0.059	-0.017	0.024	0.075	–0.509	0.237
Youngest child 4-11 yrs						
Participation rate women	0.004	-0.008	0.020	0.017	0.022	0.007
Hours worked per week women	0.173	-0.133	0.566	0.616	0.750	0.221
Participation rate men	0.000	-0.001	0.002	0.001	0.003	0.004
Hours worked per week men	0.016	0.005	-0.027	-0.001	–0.180	0.234
Model without latent classes						
Youngest child 0-3 yrs						
Participation rate women	0.017	-0.005	0.018	0.030	0.020	0.007
Hours worked per week women	0.671	-0.091	0.549	1.147	1.222	0.223
Participation rate men	0.003	-0.002	0.003	0.004	0.006	0.004
Hours worked per week men	0.069	-0.030	0.045	0.091	–0.509	0.237
Youngest child 4-11 yrs						
Participation rate women	0.002	-0.004	0.015	0.013	0.022	0.007
Hours worked per week women	0.101	-0.078	0.418	0.445	0.750	0.221
Participation rate men	0.000	-0.001	0.003	0.002	0.003	0.004
Hours worked per week men	0.020	-0.029	0.061	0.056	–0.180	0.234

[a] 2 latent classes. [b] Additional estimates on the same sample as Bettendorf, Jongen, and Muller (2012), full regression results available on request.

the elasticity is even much lower (–0.03), but this is partly the result of the lower share of women using formal care in this group.[39]

In Table 4.5 we present a test of our structural model. Bettendorf, Jongen, and Muller (2012) analyse the employment effects of the reforms discussed in Section 4.2 using difference-in-differences (DD). The identification in Bettendorf, Jongen, and Muller (2012) comes mostly from the intertemporal dimension, using a before–after comparison with data for the period 1995–2009. The identification in our analysis comes in part from intertemporal variation from the policy reforms in the period 2006–2009, but in part also from the cross-sectional variation. Bettendorf, Jongen, and Muller (2012) present estimation results for mothers with a youngest child 0–11 years of age, but this includes single mothers. Furthermore, they report effects for a different classification of mothers (with a youngest child 0–3, 4–7 and 8–11 of age). To make the comparison with the DD as clean as possible, we used the same initial sample as Bettendorf, Jongen, and Muller (2012) but we estimate responses for the subgroups we consider in our empirical analysis, that is men and women in couples with a youngest child 0–3 or 4–11 years of age. The results are given in Table 4.5, along with the simulation results for the estimated structural model.

Table 4.5 shows that the results for the structural model are very much in line with the results of the DD analysis for mothers. Indeed, we can not reject that the DD estimates for the effect on hours worked and participation of mothers are equal to the simulated effects. The estimated effects on the participation rate of fathers is again very much in line with the prediction from the structural model, and we can not reject that they are the same. For the intensive margin, for fathers with a youngest child 4–11 years of age, the DD analysis suggests a smaller negative effect on hours worked per week by the employed than the structural model, although the coefficient is not significantly different from the prediction of the structural model. The only coefficient of the DD analysis which differs significantly from the prediction of the structural model is the intensive margin response by fathers with a youngest child 0–3 years of age, for which the DD analysis suggests a larger, negative response than the structural model.

Table 4.5 also shows the predictions of the structural model when we do not allow for latent classes. In this case the predictions of the structural model move away from the DD estimates, in particular for hours worked per week by women in couples with a youngest child 4–11 years of age. Hence, a comparison with the DD analysis seems to favour a model with latent classes over a model without latent classes.

39 Table 4.A.4 in the appendix gives the resulting elasticities when we include a proxy for informal childcare. This hardly affects our results. The simulated labour supply elasticities and price elasticities of formal childcare are very similar to the model where we do not include informal care in the utility function.

4.6 RELATIVE EFFECTIVENESS OF FISCAL STIMULI

The structural model allows for a simulation of counterfactual policy reforms. Specifically, we use the model to study the effects of a number of prominent fiscal stimuli for working families with children. The policy reforms we consider are motivated by the actual reforms that have occured in the Netherlands over the past decade. Since many countries have witnessed, or are considering, similar types of reforms, we believe that the relevance of our results extends well beyond the borders of the Netherlands.

We consider three types of fiscal stimuli for working parents: i) an increase in childcare subsidies, ii) an in-work benefit for secondary earners, and iii) an in-work benefit for primary and secondary earners in couples. Furthermore, for each of these fiscal stimuli we consider an across-the-board reform, where the additional subsidy does not depend on individual or household income, and we consider a reform where the stimulus rises with income, targeted more at middle and high income households and the intensive margin of labour supply. Comparing the results for the across-the-board reforms and the income dependent reforms we can study to what extent there is an equity-efficiency trade-off for the different fiscal stimuli. In all simulations we consider the effects of a reform that costs 100 million euro[40] given the initial distribution of labour supply and childcare choices, i.e. without behavioural changes.[41] Further details of the reforms we simulate and the output we report are given below.

First we consider the following three across-the-board scenario's:

(1) An income independent increase in childcare subsidies: we increase the hourly childcare subsidy by 10.3 percentage points of the hourly price.

(2) An income independent ('flat') in-work benefit for secondary earners of 290 euro per year.

(3) An income independent ('flat') in-work benefit for primary and secondary earners in couples of 126 euro per year.

40 By doing so, we deviate from the approach chosen in Chapter 3 where we stipulated that the reform has to be revenue neutral. There, we wanted to see whether we can set up a tax-benefit system which would be more supportive of female labor supply while sustaining the same level of public spending. Here, we want to analyze how cost-effective are competing policy proposals in stimulating the female labor supply. The revenue-neutrality is of secondary importance, since neither the implemented reforms were accompanied by measures designed to compensate their budgetary burden. The revenue-neutrality of the corresponding reforms can be achieved by measures similar to the one applied in Chapter 3, but doing so should not change relative cost-effectiveness of the evaluated reforms.

41 Alternatively, we could have simulated reforms that generate the same budgetary costs after taking into account behavioural responses. However, this approach would not lead to different conclusions regarding the effectiveness of the different fiscal stimuli.

For these simulations we present the following output. First, to study the equity effects of the reforms, we present scatterplots of the redistributional effects and also report the effect on the Gini-coefficient of disposable income. Next, we report the effects on labour supply, both in terms of participation (extensive margin) and in total hours worked (extensive and intensive margin). Next to the effect on labour supply and labour production we also report the effect on formal childcare use. Finally, we measure the relative effectiveness of the different fiscal stimuli by looking at public spending required to bring an additional person to the workforce.

The redistributional effects of the across-the-board reforms are given in Figure 4.5, 4.7 and 4.9, respectively. On the horizontal axis we have disposable household income, on the vertical axis the percentage change in disposable household income relative to the base (in %). None of these subsidies do depend on income directly, the absolute change in disposable income is the same for low and high income households that use the same formal childcare and have the same number of partners employed. However, because we present redistributional effects in percentage terms, percentage changes in disposable income are lower for high income households, ceteris paribus. For the childcare reform in Figure 4.5 we see a number of 'lines', as families differ in the number of children they have, as well as the number of days of formal childcare they use per child. Furthermore, for the childcare reform we further observe that although the change in the subsidy itself does not depend on income, there is quite a large mass of families with middle and higher incomes that receive susbtantially more subsidy. This is because these families are more likely to use formal childcare than lower-income families. Indeed, the effect of the reform on the Gini-coefficient (without behavioural changes), reported in column (1) in Table 4.6, is actually positive for the across-the-board childcare reform (inequality rises). Figure 4.7 shows two lines for the in-work tax credit of secondary earners. A first line lies on the x-axis and represents couples who do not receive the in-work tax credit since at least one of the partners does not work. The second line refers to couples in which both partners work and therefore receice the in-work tax credit. Looking at the percentage changes, we see that a large part of the credit goes to households with a lower income. Indeed, for this reform we see a modest decline in the Gini-coefficient for disposable incomes, see column (2) in Table 4.6. Finally, Figure 4.9 gives the redistributional effects of the in-work benefit for both primary and secondary earners. Here we see three lines, one for two-earner households, one for one-earner households and one for households in which neither of the two parents works. Looking at the percentage changes, this reform is targeted even more at lower incomes, and the Gini-coefficient falls the most in this scenario, see col. (3) in Table 4.6.

Effect on initial incomes, childcare subsidies:

Figure 4.5: Across-the-board

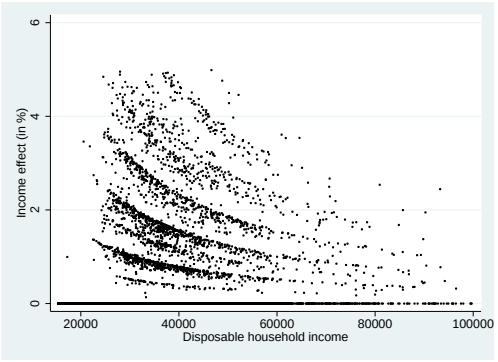
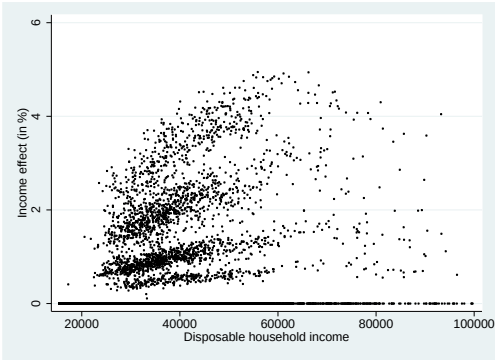


Figure 4.6: Targeted at higher incomes



Effect on init. incomes, in-work benefit for secondary earners with children:

Figure 4.7: Across-the-board

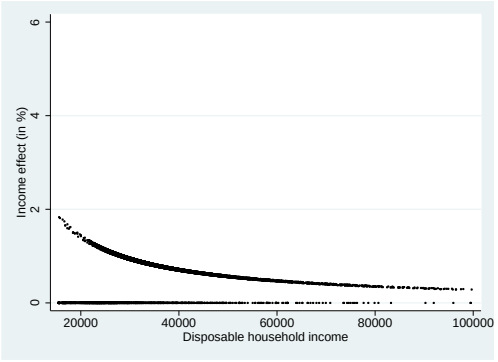
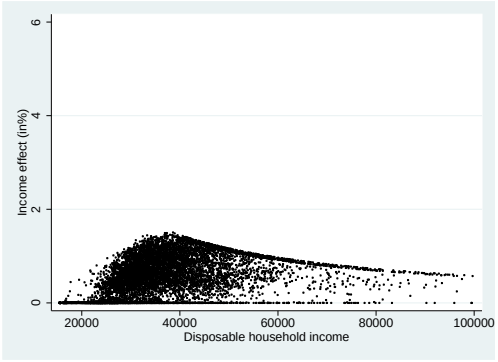


Figure 4.8: Targeted at higher incomes



Effect on initial incomes, in-work benefit for all workers with children:

Figure 4.9: Across-the-board

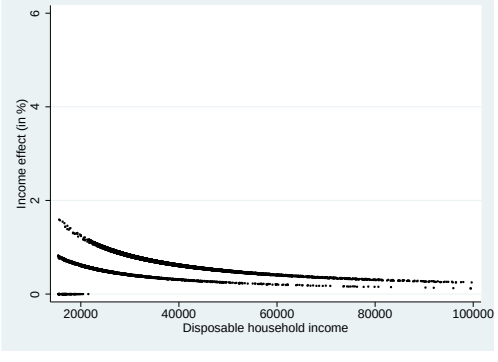


Figure 4.10: Targeted at higher incomes

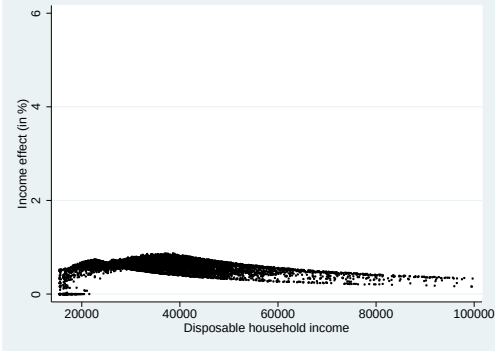


Table 4.6 presents the effects on labour participation, formal childcare and government finances of these three scenario's. Column (1) gives the results for the increase in the childcare subsidy. First, consider the effects on labour participation of couples with a youngest child 0–3 years of age (pre-primary school). The higher childcare subsidy draws some men into the workforce (extensive margin). But men already working reduce their hours worked (intensive margin).[42]

42 The intensive margin also captures a composition effect, when new entrants on the labour market work different hours on average than the incumbent workforce.

Table 4.6: Effectiveness of fiscal stimuli of 100 million euro

	Across-the-board			Targeted more at higher incomes		
	(1)	(2)	(3)	(4)	(5)	(6)
	Childcare subsidies	I-WB 2^{nd} earners	I-WB all	Childcare subsidies	I-WB 2^{nd} earners	I-WB all
	Percentage changes					
Gini coefficient	0.35	-0.10	-0.34	0.93	0.53	-0.14
Labour supply total	0.55	0.28	0.09	0.56	0.44	0.11
Labour supply youngest child 0–3						
– Men	0.04	0.03	–0.02	0.10	0.02	0.02
— Extensive margin	0.22	0.12	0.05	0.16	0.11	0.06
— Intensive margin	–0.19	–0.09	–0.07	–0.06	–0.09	–0.04
– Women	2.44	0.75	0.23	2.29	1.34	0.27
— Extensive margin	1.52	0.89	0.34	1.19	0.73	0.13
— Intensive margin	0.90	–0.14	–0.11	1.10	0.61	0.14
Labour supply youngest child 4–11						
– Men	0.03	–0.02	–0.03	0.05	–0.03	–0.01
— Extensive margin	0.02	0.07	0.03	0.02	0.06	0.04
— Intensive margin	0.01	–0.08	–0.06	0.03	–0.09	–0.05
– Women	0.89	1.02	0.44	0.91	1.49	0.41
— Extensive margin	0.39	1.24	0.56	0.35	0.80	0.23
— Intensive margin	0.50	–0.21	–0.12	0.55	0.68	0.18
Formal childcare total	12.62	1.28	0.67	11.13	2.12	0.79
Formal childcare youngest child 0–3	11.54	1.20	0.64	9.51	1.82	0.70
Formal childcare youngest child 4–11	16.28	1.54	0.80	16.61	3.12	1.09
	Millions of euro					
Additional public exp. ex ante	100.0	100.0	100.0	100.0	100.0	100.0
Knock-on effect childcare subsidies	132.4	11.0	5.8	103.4	16.8	6.3
Knock-on effect taxes and benefits	–52.7	–19.7	–4.8	–60.1	–33.5	–8.5
Additional public exp. ex post	179.7	91.3	101.0	143.3	83.3	97.8
	Euro					
Ex ante spending per FTE	28,135	55,269	179,070	27,782	35,211	142,829
Incl. effect on formal childcare use	65,374	61,323	189,402	56,509	41,122	151,859
Incl. effect on taxes and benefits	50,559	50,442	180,772	39,810	29,328	139,741

The overall effect on hours worked by men is small. The effect on hours worked by women is much more pronounced. Indeed, there is a substantial positive effect on both the extensive and the intensive margin. Here, it is important to note that although the increase in the childcare subsidy is income independent, there is a substitution effect at work. Indeed, mothers that work more hours are also more likely to use more formal childcare, and thus have an extra benefit from the increase in the childcare subsidy. Turning to couples with a youngest child 4–11 years of age, we observe similar though somewhat smaller labour supply effects. Children in primary school are less likely to go to formal childcare, and if they do they typically go for only a few hours per day. For all couples with a child 0–11 years old, we find an increase in hours worked of 0.55%.

In absolute terms, men in couples increase their labour supply by 0.01 hours per week. The average increase is much higher for women in couples: 0.44 and 0.13 for mothers with a youngest child 0–3 and 4–11 years of age, respectively. However, the rise in formal childcare is much more pronounced than the rise in total hours worked. Couples with a youngest child 0–3 (4–11) years of age demand 1.77 (0.33) additional hours of childcare per week. This reflects the fact that there is not a 1-to-1 link with hours worked.

To conclude the analysis of reform (1), at the bottom of the table we give the effects on government finances, excluding and including so-called knock-on effects on government finances that result from behavioural changes. In case of the childcare subisdy reform, the increase in hours worked increases tax receipts and reduces expenditures on (welfare) benefits. This shift is however dominated by the increase in childcare subsidies due to substitution of other types of care for formal childcare. The average subsidy rate for formal childcare in the baseline scenario is 76% of the hourly price, making the increase in formal childcare rather costly to the government. Taking into account the behavioural responses, government expenditures rise by 180 instead of 100 million euro. As a measure of the relative effectiveness of reform (1), in the last three rows we calculate the additional public spending per additional fulltime-worker equivalent (fte) employed. Ignoring the knock-on effects, additional public spending per additional fte is 28 thousand euro. However, when we take into account the increase in formal childcare, additional public spending per additional fte rises to 65 thousand euro. Finally, taking into account the savings on benefits and the additional tax receipts, we still arrive at 51 thousand euro per additional fte. These calculations highlight that it is important to take into account the knock-on effects of changes in formal childcare when calculating the effectiveness of childcare subsidies.

Column (2) in Table 4.6 gives the behavioural responses and corresponding budgetary effects for the 'flat' in-work benefit targeted at secondary earners. First, again consider the effects on couples with

a youngest child 0–3 years of age. The effect on the labour supply of men is again small, most men are not secondary earners. The effect on hours worked by women is much smaller compared to reform (1). In part, this is simply due to the fact that a larger part of the tax credit actually goes to the larger group of mothers with a youngest child 4–11 years of age than with reform (1). However, we also see that for reform (2) the intensive margin response is negative for women with a youngest child 0–3 years of age, because the in-work benefit only has an income effect on the intensive margin. Turning to the couples with a youngest child 4–11 years of age, we find a larger effect on hours worked compared to reform (1), as a larger part of the subsidies goes this group. Also for this group, the effect on the intensive margin is negative. When we look over all couples, we find that the increase in total hours worked in reform (2) is only about half of reform (1).

Since the tax credit does not affect the price of formal childcare for parents, reform (2) has only a modest effect on the use of formal childcare, following the increase in total hours worked. This is also reflected in the knock-on effects. As the flat tax credit for secondary earners is less succesful in raising hours worked, the knock-on effect of increases in taxes and savings on benefits is smaller than in the childcare subsidy reform. However, because the knock-on effect on childcare subsidy expenditures is much smaller than reform (1), reform (2) generates a positive knock-on effect of 9 million euro. Again, we calculate the relative effectiveness of the reform as additional public expenditures per additional fte employed. Ignoring the knock-on effects, additional public spending per additional fte employed is 55 thousand euro. Taking into account the knock-effects, this number becomes 50 thousand euro. When we compare the relative effectiveness of reform (2) with reform (1), we find that both reforms are about equally effective. Note however, that reform (1) goes at the expense of greater inequality, whereas reform (2) actually reduces inequality.

Finally, column (3) in Table 4.6 gives the results of a flat in-work benefit for both primary and secondary earners. In this senario, a large part of the subsidies goes to the men in couples with children, who hardly respond to financial incentives. As a result, the effects are much smaller than in reform (1) and (2). We still see the positive effect on the extensive margin, and the negative effect on the intensive margin (due to the income effect). The increase in total hours worked is just 0.09%. The knock-on effects are therefore also small, and close to zero overall. This makes this the most expensive reform in terms of additional spending per additional hour worked. Indeed, additional public spending per additional fte employed is in the order of 180 thousand euro, making this policy relatively ineffective when compared to reform (1) and (2). We should note though, that this reform leads to a bigger drop in inequality than reform (2), and does not raise inequality like reform (1). With this effect on equity in mind, the re-

sults for reforms (1)–(3) suggest that the shift in the Netherlands from in-work benefits for all parents to in-work benefits just for secondary earners and to higher childcare subsidies was an effective one in terms of raising total hours worked.

Next, we consider the effectiveness of fiscal stimuli when the fiscal stimuli are targeted more at middle and higher incomes and the intensive margin. Specifically, we study the effects of the following three scenario's:

(4) An income dependent increase in childcare subsidies: we increase the hourly childcare subsidy so that the parental fee falls by 41% for all incomes. Given that middle and higher incomes pay a larger fee in the baseline, this reform targets mostly middle and high income families.

(5) An income dependent in-work benefit for secondary earners starting at 4,000 euro, and then rising with 2.2% of taxable income of the secondary earner up to a maximum of 581 euro per year at an individual income of 30,000 euro.

(6) An income dependent in-work benefit for primary and secondary earners in couples, starting at 4,000 euro, and then rising with 0.6% of taxable income of the primary or secondary earner up to a maximum of 168 euro per year at an individual income of 30,000 euro.

The redistributional effects of these reforms are given in Figure 4.6, 4.8 and 4.10, respectively. The subsidies now rise with income, and we see that the percentage changes in disposable income are typically smaller for lower incomes and bigger for higher incomes when compared to the reforms considered in Figure 4.5, 4.7 and 4.9. Indeed, reforms (4) and (5) increase income inequality, as measured by the Gini-coefficient, more than reforms (1) and (2), see Table 4.6. Furthermore, reform (6) reduces inequality less than reform (3). When there is a trade-off between equity and efficiency, we expect these reforms to be more effective in terms of labour supply and public spending per additional fte employed. But is this actually true, and if so, how much of a difference does it make?

Column (4) in Table 4.6 gives the effects of the increase in childcare subsidies targeted more at middle and high incomes. Starting with the couples with a youngest child 0–3 years of age, the effect on hours worked by men is still limited, though somewhat larger than reform (1), as the rise in the subsidy with income mitigates the negative intensive margin effect. The effect on hours worked by women is actually smaller than in reform (1). The intensive margin response is bigger, but this is dominated by a smaller effect on the extensive margin. The results for couples with a youngest child 4–11 in reform (4) are not that different from reform (1), though the effect on the extensive

and intensive margin for women is somewhat smaller respectively larger. The effect on overall hours worked is quite similar in reform (4) when compared to reform (1). Hence, targeting childcare subsidies on middle and higher incomes is not necessarily better when looking at total hours worked. This suggests that the current system in the Netherlands which targets subsidies mostly at low incomes makes sense when it comes to hours worked. Furthermore, the increase in formal childcare is somewhat smaller in reform (4) than in reform (1), the reform leads to less substitution of other types of care for formal childcare for couples with a youngest child 0–3 years of age.

The knock-on effects are more favourable for reform (4) than reform (1). The additional hours worked by middle and higher incomes generate more additional tax revenue per additional hour worked. Furthermore, substitution of other types of care for formal care is less costly for the government, as the subsidy per hour of formal childcare is lower for middle and higher incomes than for lower incomes. With an about equal effect on total hours worked and much more favourable knock-on effects, it comes as no surprise that additional public spending per additional fte employed is more favourable in reform (4) than in reform (1), with 40 thousand euro in reform (4) compared to 51 thousand euro in reform (1). However, the difference comes at the expense of additional income inequality, and hence, once we take into account the knock-on effects on the government budget, there is actually a trade-off between equity and efficiency when it comes to the targeting childcare subsidies.

Column (5) gives the results for the in-work benefit for secondary earners that rises with income. Compared to the flat in-work benefit for secondary earners, reform (2), this reform has a more favourable effect on hours worked of women. Indeed, the substitution effect of this reform makes the intensive margin responses by women positive rather than negative. The effect on total hours worked is also considerably larger than reform (2), although still smaller than the childcare reforms (1) and (4). However, because this reform does not generate a large response in the use of formal childcare, the knock-on effects are rather favourable: plus 17 million euro. When we calculate the additional expenditures per additional fte employed, reform (5) is the most effective, with 29 thousand euro per additional fte employed. This suggests that the Dutch reform in 2009, making the in-work benefit for secondary earners income dependent, was rather effective. However, also here there is a trade-off with equity, as the additional hours worked come at the expense of additional income inequality.

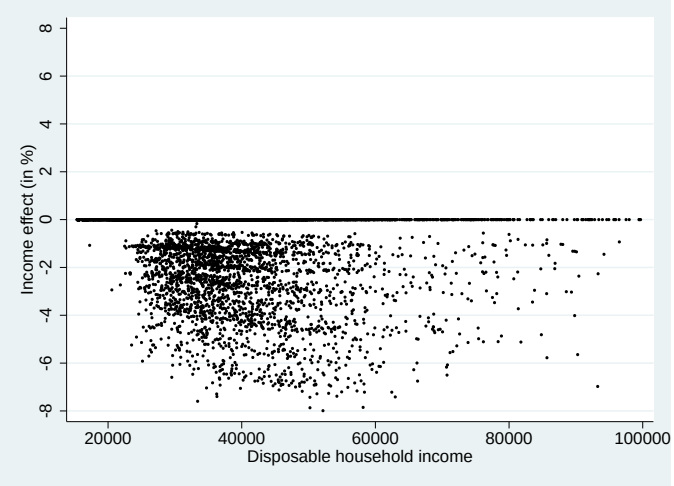
Finally, column (6) gives the results of the income dependent tax credit for primary and secondary earners. The overall effect on hours worked and government finances is slightly better than for the flat credit for primary and secondary earners. Again, there is a trade-off between efficiency and equity. However, this reform still has only a

marginal effect on overall hours worked, and with 140 thousand euro per additional fte employed is still rather expensive.

4.7 SIMULATING THE 2011-2013 CHILDCARE REFORM

The empirical model can also be used to simulate the effects of recent cuts in childcare subsidies which were taken by the Dutch government. Following the steep rise in public expenditures on formal childcare over the period 2005–2009, and after the Dutch economy was hit by the Great Recession, the Dutch government announced to cut expenditures on childcare over the period 2011–2013. As a result, the average contribution rate of households to formal childcare was projected to increase from 22% to 34% (Ministry of Social Affairs and Employment, 2011). The redistributional effects on disposable household income are shown in Figure 4.11. The simulated effects on labour participation, formal childcare use and government finances are given in Table 4.7.

Figure 4.11: Redistributional effects childcare reform 2011–2013



The reform is projected to have only a small negative effect on hours worked by fathers. The effect is more pronounced for mothers, in particular for mothers with a youngest child 0–3 years of age. Their hours worked drop by 3.4%, of which a substantial part is on the intensive margin. The drop in the use of formal childcare is projected to be much bigger in percentage terms, 14% respectively 20% for households with a youngest child 0–3 years and 4–11 years of age. As a result, the knock-on effect for the government budget is actually positive. Additional savings on childcare subsidies more than offset the loss in tax receipts and the rise in benefit expenditures. We should also note that the predicted decline in the use of formal childcare is actually quite similar to what is observed following the recent cuts

Table 4.7: Simulation results: childcare reform 2011–2013

	Perc. changes	Perc. changes
	Youngest child 0–3 yrs	Youngest child 4-11 yrs
Labour supply men	–0.15	–0.05
– Extensive margin	–0.26	–0.02
– Intensive margin	0.11	–0.03
Labour supply women	–3.43	–1.12
– Extensive margin	–1.88	–0.45
– Intensive margin	–1.58	–0.67
Formal childcare	–14.24	–19.87

Overall effect

Perc. changes			Millions of euro
Gini coefficient	–1.05	Additional public exp. ex ante[a]	–154.4
Labour supply total	–0.79	Knock–on effect childcare subsidies	–109.2
Formal childcare total	–15.53	Knock-on effect taxes and benefits	82.8
		Additional public exp. ex post[a]	–180.8

[a] Additional public expenditures in our sample.

in childcare subsidies, with the use of formal childcare falling by 18% (Ministry of Social Affairs and Employment, 2014). However, uncertainty about trend growth absent the reform, and the effect of the Great Recession, complicate the comparison.

4.8 CONCLUSION

We have estimated a structural model for couples with a youngest child 0–3 and 4–11 years of age, where we model the simultaneous choice over hours worked by fathers, mothers and the hours of formal childcare use. Large exogenous variation in childcare subsidies and in-work benefits for working parents benefits the identification of the structural parameters. Furthermore, we account for unobserved heterogeneity by using a flexible framework of latent class models. The model produces labour supply responses to reforms over the period 2005–2009 quite similar to a DD analysis on the same reforms performed by Bettendorf, Jongen, and Muller (2012). The model also predicts the steep decline in the use of formal childcare observed following the recent cuts in childcare subsidies.

We use this model to study the relative effectiveness of different types of fiscal stimuli for working parents with children 0–11 years of age. We find that an across-the-board increase in childcare subsidies is more effective than an across-the-board increase in in-work benefits for this group of working parents. That is, when we ignore the knock-on effects on the government budget due to behavioural responses.

However, because childcare subsidies also lead to a shift from other types of childcare to formal childcare, the knock-on effect on the government budget is actually negative, despite higher tax receipts and savings on benefits. Comparing the additional public expenditures per additional hour worked, an across-the-board increase in childcare subsidies is about equally effective as an across-the-board increase of in-work benefits for secondary earners. Both policies are much more effective than an across-the-board increase of in-work benefits for both primary and secondary earners, as labour supply by primary earners is rather unresponsive to financial incentives. These results support the move from the Dutch government over the period 2005–2009 to abolish the joint in-work benefit for both primary and secondary earners, and increase childcare subsidies and the in-work benefit for secondary earners alone.

We also consider the extent to which childcare subsidies and in-work benefits targeted more at middle and higher incomes are more effective in raising total hours worked than an across-the-board increase in these policies. Indeed, we want to quantify the trade-off between efficiency and equity. Our model shows that targeting childcare subsidies more at middle and high income families has almost the same effect on total hours worked as an across-the-board increase in childcare subsidies, but of course leads to a less equitable income distribution. This finding can motivate the current setup of the Dutch system, where low income families receive higher childcare subsidies than middle and high income families. However, targeting childcare subsidies more at middle and high income families does increase hours worked by a relatively productive group, which results in higher knock-on effects of taxes. Furthermore, as middle and high incomes receive less subsidy per hour of formal childcare, substitution of other types of care for formal care is less costly for the government. Both factors lead to lower additional public expenditures per additional hour worked for childcare subsidies targeted more at middle and higher incomes, once we take into account the behavioural effects on the government budget. So in the end, there is a trade-off between equity and efficiency when it comes to childcare subsidies.

An in-work benefit for secondary earners that rises with income is substantially more effective in terms of total hours worked and government finances than a 'flat' in-work benefit for secondary earners. So, from an efficiency point of view, introducing the income dependent part to the in-work benefit for secondary earners in 2009 by the Dutch government made sense. However, this reform came at expense of higher income inequality, and here we clearly face an equity-efficiency trade-off.

There are still a number of limitations to our current analysis that we would like to overcome in future work. An interesting next step would be to model these decisions in a life cycle model (Blundell, Bozio, and

Laroque, 2013). Indeed, there may be career effects extending beyond the period when the children are young. Another interesting avenue to consider is the effect of participation in formal childcare on the well-being of children and how they fare later in life, and whether or not there is a difference between children from low income and high income families (Havnes and Mogstad, 2014). Finally, the childcare reform may have been more salient than the reform of in-work benefits. Indeed, Chetty, Looney, and Kroft (2009) stress the importance of salience in the behavioural responses to taxes and subsidies, and this too seems an interesting topic for future research.

4.9 ACKNOWLEDGMENTS

We have benefitted from comments and suggestions by Rolf Aaberge, Hans Bloemen, Tom Kornstad, Andrew Shephard, Arthur van Soest, Thor Thoresen, Trine Vatto, participants of the CPB Workshop Behavioural Responses to Taxation and Optimal Tax Policy May 2013 in The Hague, the IIPF Conference August 2013 in Taormina, the ESEM European Meetings August 2013 in Gothenburg, the ZEW Family Economics Workshop November 2013 in Mannheim, the SOLE Annual Meetings May 2014 in Arlington and seminar participants at CPB Netherlands Bureau for Economic Policy Analysis and Tilburg University. Remaining errors are our own. We are grateful to Statistics Netherlands for constructing the data set, and to the Ministry of Social Affairs and Employment for co-financing the construction of the data set.

4.A WAGE EQUATIONS

For the employed we use observed wages. For the non-employed we simulate wages. To this end, we run wage regressions by sex and then by level of education, where education is split into three levels (lower, middle and higher educated).

We use panel data techniques to account for unobserved individual-specific effects. We performed a Hausman test in order to test whether random effects or fixed effects are appropriate. For all groups, we reject the null hypothesis that the individual-specific effects are uncorrelated with regressor and therefore we prefer fixed effects over random effects estimation. However, we lose information on time-invariant regressors with fixed effects and therefore opt for the quasi-fixed effects model (Mundlak, 1978).

To account for the possibility of selection we first estimate the probability of participation using a pooled probit regression

$$p_{it} = x'_{it}\gamma + z'_{it}\theta + \nu_{it}, \tag{4.A.1}$$

where vector z_{it} contains variables that are expected to have an effect on the probability of participation but not on wages (an exclusion restriction). From this regression we determine the inverse Mills' ratio

$$invmills_{it} = \phi(p_{it})/\Phi(p_{it}). \tag{4.A.2}$$

The inverse Mills' ratio is then included in the quasi-fixed effects model

$$ln(w_{it}) = x'_{it}\beta + \omega_i + \bar{x_i}'\pi + \lambda_t invmills_{it} + \epsilon_{it} \tag{4.A.3}$$

where the individual specific effect consists of a random part, ω_i with $\sim IID(0, \sigma^2_\omega)$, and a part which is allowed to be correlated with regressors $\bar{x_i}'\pi$. Here, $\bar{x_i}$ is the average of time-varying variables such as age. A significant coefficient for an element of π provides evidence that the individual specific effect is correlated with one of the regressors.

Table 4.A.1 shows estimation results for all subgroups. We use age splines since we expect that the relationship between wage and age is nonlinear. Table 4.A.1 shows that age increases with age but at a diminishing rate. This is in line with other studies (Vella and Verbeek, 1998, 1999). For both singles and couples we see that the age profile is steeper for higher educated individuals. We also include cohort and year dummies in the regression. Because of perfect collinearity between age, cohort and period we use transformed time dummies following Deaton and Paxson (1994). The time dummies for 2006 and 2007 depend on the dummies for later years and are calculated manually.[43] Year dummies are significant in most specifications while the cohort variables are jointly significant for most subgroups. Wages

43 t2006=-(d2007+d2008+d2009) and t2007=-2*d2008-3*d2009

Table 4.A.1: Wage equations

	Men			Women		
	Lower educ.	Middle educ.	Higher educ.	Lower educ.	Middle educ.	Higher educ.
Age						
18–30	0.045***	0.047***	0.076***	0.037***	0.037***	0.047***
31–40	0.020***	0.029***	0.045***	0.022***	0.024***	0.035***
41–50	0.013***	0.020***	0.028***	0.024***	0.021***	0.023***
51–63	0.010***	0.008***	0.011***	0.020***	0.017***	0.013***
Cohort[a]						
1980–1989	0.085***	0.147***	0.173***	0.146***	0.126***	0.158***
1975–1980	0.025	0.074***	0.129***	0.063***	0.080***	0.118***
1970–1975	0.019*	0.034***	0.093***	0.030***	0.048***	0.077***
1960–1965	0.010	−0.017***	−0.012	−0.008	−0.019***	−0.044***
1955–1960	−0.002	−0.031***	−0.043***	0.009	−0.027**	−0.064***
<1955	0.007	0.002	−0.012	0.010	−0.019*	−0.046***
Ethnicitity[a]						
Western immigrant	0.003	−0.068***	−0.055***	0.001	−0.026***	−0.032***
Non-Western immigrant	−0.062***	−0.231***	−0.291***	−0.051***	−0.074***	−0.114***
Partner						
Married	0.015***	0.017***	0.015***	−0.011**	−0.015***	−0.025***
Year						
2006	0.005	0.005	0.004	0.006	0.004	0.002
2007	−0.006	−0.006	−0.003	−0.007	−0.005	−0.002
2008	−0.002***	−0.003***	−0.007***	−0.004***	−0.003***	−0.003***
2009	0.004***	0.004***	0.005***	0.005***	0.004***	0.002***
Mundlak averages age						
18–30	−0.008*	0.000	−0.005	−0.003	−0.002	0.001
31–40	−0.006**	−0.003**	0.000	−0.012***	−0.008***	−0.004***
41–50	−0.008***	−0.007***	−0.014***	−0.022***	−0.016***	−0.017***
51–63	−0.008***	−0.015***	−0.019***	−0.018***	−0.020***	−0.019***
Inverse Mills' ratio	−0.329***	0.452***	0.674***	−0.008	0.026**	0.098***
Attrition indicator	−0.004	−0.001	−0.001	−0.004	−0.004	0.000
Constant	1.446***	1.162***	0.618***	1.298***	1.430***	1.273***
Observations	88,997	168,316	129,663	60,824	146,294	89,859
Number of individuals	26,779	49,634	37,742	19,385	44,262	26,770

[a] Reference group: born in 1965–1970 and autochtonous.

are lower on average for non-Western immigrants. The coefficients for the Mundlak age averages are joinly significant in all specifications, but have no economic interpretation.

The lower part of Table 4.A.1 shows that the inverse Mills' ratio is significant for most groups. Hence, we have evidence that selection bias is present for most groups. We also include an attrition indicator in order to test for the presence of attrition bias.[44] The attrition indicator is not significant for all subgroups.

4.B PRICE EQUATIONS FORMAL CHILDCARE

For non-users of formal childcare we have to simulate a price for childcare. We have information on the use of formal childcare in the Netherlands for the period 2006–2009. Here, a distinction is made between daycare (children 0–3 years of age) and out-of-school care (children 4–11 years of age).

Again, we estimate a quasi-fixed effects model for the prices of daycare and out-of-school care.[45] Here, we follow the same procedure as for the wage estimations and estimate the following price equation:

$$p_{it} = x'_{it}\beta + \omega_i + \bar{x_i}'\pi + \lambda_t invmills_{it} + \epsilon_{it} \tag{4.A.4}$$

where the individual specific effect consists of a random part, ω_i with $\sim IID(0, \sigma^2_\omega)$, and a part which is allowed to be correlated with regressors $\bar{x_i}'\pi$. Here, $\bar{x_i}$ is the average of age which does not vary over time. Our dependent variable is the hourly real price.

We focus on households since childcare is consumed at the household level. As it turns out, characteristics of females are more important in predicting use and price of childcare than characteristics of men. Hence, we only include females characteristics in the regressions.

Table 4.A.2 shows estimation results for daycare and out-of-school care.[46] Estimation results show that year dummies are significantly increasing for daycare. However, time effects are less important in the price equation for out-of-school care. Households with higher educated women or younger women pay a higher price on average. We do not find evidence that selection bias or attrition bias are present.

44 The attrition indicator is a dummy which equals 1 if an individual leaves the sample in our data period 2006-2009.

45 We conduct a Hausman test in order to test whether fixed or random effects is appropriate. In all cases, the Hausman test favours the fixed effects model.

46 Including a squared term for age, age splines, ethnicity, a dummy for age of the youngest child or a dummy for multiple children one at a time, leads to insignificant coefficients for each of these variables.

Table 4.A.2: Price equation formal childcare

	Daycare	Out-of-school care
Year		
2007	0.058***	0.015
2008	0.123***	0.025
2009	0.153***	0.035
Higher educated women	0.000	0.020*
Age women	–0.017***	-0.031***
Single parent	0.033**	-0.047***
Mundlak age average	0.014**	0.026**
Inverse Mills' ratio	–0.032	-0.008
Attrition indicator	–0.001	0.005
Constant	5.507***	5.741***
Observations	35,675	28,938
Households	14,984	12,015

4.C ELASTICITIES AND SHARES WITH NEGATIVE MARGINAL UTILITY BY NUMBER OF LATENT CLASSES

Table 4.A.3: Elasticities by number of latent classes

youngest child 0–3 yrs	1 LC	2 LC	3 LC	4 LC
Gross hourly wage men +1%				
Labour supply men	0.09	0.08	0.03	0.09
– Extensive margin	0.08	0.07	0.01	0.07
– Intensive margin	0.01	0.01	0.02	0.02
Labour supply women	–0.15	–0.15	–0.21	–0.15
Formal childcare	0.10	0.11	0.05	0.05
Gross hourly wage women +1%				
Labour supply women	0.37	0.40	0.33	0.48
– Extensive margin	0.25	0.25	0.18	0.30
– Intensive margin	0.12	0.15	0.15	0.18
Labour supply men	–0.04	–0.05	–0.07	–0.06
Formal childcare	0.40	0.41	0.44	0.45
Gross price formal childcare +1%				
Formal Childcare	–0.61	–0.66	–1.09	–0.92
Labour supply men	0.00	0.00	0.00	0.00
Labour supply women	–0.13	–0.14	–0.15	–0.16
Observed choices with negative marginal utility income	0.00	0.00	0.00	0.00
Observed choices with negative marginal utility leisure men	0.74	0.39	0.00	0.08
Observed choices with negative marginal utility leisure women	0.38	0.36	0.09	0.00
Observed choices with negative marginal utility formal childcare	0.54	0.35	0.59	0.49
youngest child 4–11 yrs	1 LC	2 LC	3 LC	4 LC
Gross hourly wage men +1%				
Labour supply men	0.09	0.06	0.08	0.08
– Extensive margin	0.08	0.04	0.07	0.06
– Intensive margin	0.01	0.02	0.01	0.02
Labour supply women	–0.11	–0.07	–0.10	–0.11
Formal childcare	0.27	0.15	0.24	0.22
Gross hourly wage women +1%				
Labour supply women	0.38	0.47	0.44	0.48
– Extensive margin	0.25	0.31	0.31	0.29
– Intensive margin	0.13	0.16	0.13	0.19
Labour supply men	–0.03	–0.04	–0.07	–0.05
Formal childcare	0.45	0.77	0.71	0.83
Gross hourly price formal childcare +1%				
Formal Childcare	–0.36	–0.77	–0.70	–0.83
Labour supply men	0.00	0.00	0.00	0.00
Labour supply women	–0.02	–0.04	–0.04	–0.05
Observed choices with negative marginal utility income	0.00	0.00	0.09	0.08
Observed choices with negative marginal utility leisure men	0.78	0.17	0.26	0.41
Observed choices with negative marginal utility leisure women	0.41	0.22	0.34	0.02
Observed choices with negative marginal utility formal childcare	0.57	0.16	0.10	0.57

4.D ROBUSTNESS CHECK: INCLUDING PROXY FOR INFORMAL CHILDCARE

Table 4.A.4: Elasticities for models w/o and w/ proxy informal childcare

	Couples 0-3 yrs		Couples 4-11 yrs	
	1 LC	2 LC	1 LC	2 LC
Model without proxy informal care				
Labour supply elasticity men	0.09	0.08	0.09	0.06
Labour supply elasticity women	0.37	0.40	0.38	0.47
Price elasticity formal childcare	–0.61	–0.66	-0.36	-0.77
Model with proxy informal care				
Labour supply elasticity men	0.09	0.07	0.10	0.06
Labour supply elasticity women	0.37	0.41	0.41	0.48
Price elasticity formal childcare	–0.62	-0.70	-0.42	-0.84

4.E PREFERENCES AND FIT OF PREFERRED MODEL

Table 4.A.5: Preferences by latent class, youngest child 0–3 yrs

Latent class	1	2		1	2
Income	6.164**	15.812***	Fixed costs men	-8.885***	-11.758***
Leisure men	-66.223***	-74.155***	*Lower education	1.539**	0.522
*Age	0.367	0.663	*Middle education	1.483***	1.124
*Age^2	0.260	-1.393	*Non-Western immigrant	-0.830	-0.558
			*Western immigrant	-1.682***	-1.125*
Leisure female	-21.914***	-19.814**			
*Age	2.936	1.375	Fixed costs women	-2.520***	-2.550***
*Age^2	2.348	2.872	*Lower education	0.836	-0.674**
			*Middle education	0.484*	0.162
Income^2	2.250**	-3.646***	*Non-Western immigrant	-1.144***	-1.412***
Income*leisure men	21.444***	-2.799	*Western immigrant	-0.284	-0.868**
Income*leisure women	5.391	-8.189			
Leisure men^2	-48.270	-14.755***	Fixed cost childcare	0.690	0.365
Leisure women^2	-126.255***	-167.628***	*Non-Western immigr. men	-0.254	-0.466
Leisure men*leisure women	-0.392	-11.813	*Western immigrant men	0.993	-0.664
			*Lower education men	-0.428	-0.287
Childcare	-2.895***	-1.637**	*Middle education men	-0.267	-0.477**
*Urban area	0.643**	0.992***	*Non-Western immigr. women	-1.598	-1.261
*Non-Western immigr. men	-0.644	-0.135	*Western immigrant women	-0.999	-0.147
*Western immigrant men	0.841	0.587	*Lower education women	-1.737***	-0.766**
*Non-Western immigr. women	0.999	0.979	*Middle education women	-0.461**	-0.652***
*Western immigrant women	0.365	0.164	*Urban area	-0.859	-1.619**
Childcare^2	0.878	-0.135	Relative class shares	48%	52%
Childcare*income	0.943***	0.477*			
Childcare*leisure men	0.854	1.159			
Childcare*leisure women	-5.781***	-7.935***			

*** p<0.01, ** p<0.05, * p<0.1.

Table 4.A.6: Preferences by latent class, youngest child 4–11 yrs

Latent class	1	2		1	2
Income	3.216	3.187***	Fixed costs men	1.475***	1.464***
Leisure men	13.879	14.486***	*Lower education	0.527	0.463
*Age	2.782	1.866***	*Middle education	0.630*	0.417
*Age^2	1.246**	1.216***	*Non-Western immigrant	0.622	0.484***
			*Western immigrant	0.622*	0.636***
Leisure female	7.477	7.318***			
*Age	1.362	1.428*	Fixed costs women	0.279***	0.287***
*Age^2	1.722	1.518***	*Lower education	0.271**	0.269
			*Middle education	0.232	0.235
Income^2	1.115	1.323***	*Non-Western immigrant	0.282	0.287***
Income*leisure men	5.684	6.258***	*Western immigrant	0.298	0.284
Income*leisure women	4.855	5.615***			
Leisure men^2	28.321	29.344***	Fixed cost childcare	0.469***	0.455***
Leisure women^2	12.140***	12.445***	*Non-Western immigr. men	1.863	178.637
Leisure men*leisure women	15.919	16.436	*Western immigrant men	1.226	2.655
			*Lower education men	0.378	0.357
Childcare	1.780**	1.150***	*Middle education men	0.276*	0.287
*Urban area	0.435*	0.591	*Non-Western immigr. women	0.873	0.869
*Non-Western immigr. men	1.183*	8.534	*Western immigrant women	0.696**	0.751
*Western immigrant men	0.860	1.366	*Lower education women	0.564***	0.399*
*Non-Western immigr. women	0.725	0.640	*Middle education women	0.266	0.279
*Western immigrant women	0.556***	0.823	*Urban area	0.544	0.569*
Childcare^2	0.299	0.346***	Relative class shares	42%	58%
Childcare*income	0.385	0.417***			
Childcare*leisure men	2.713***	2.860*			
Childcare*leisure women	1.495***	1.639***			

*** p<0.01, ** p<0.05, * p<0.1.

Fit labour supply men

Figure 4.A.1: Age youngest child 0–3 yrs

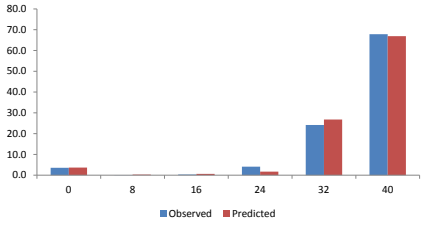
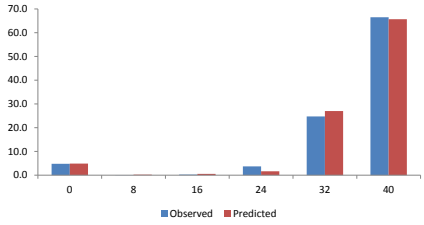


Figure 4.A.2: Age youngest child 4–11 yrs



Fit labour supply women

Figure 4.A.3: Age youngest child 0–3 yrs

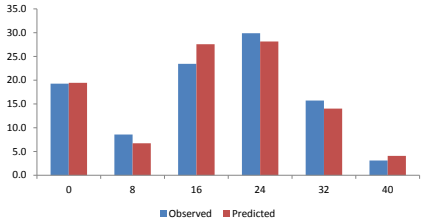
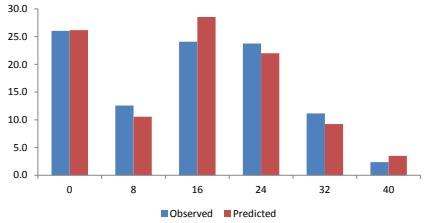


Figure 4.A.4: Age youngest child 4–11 yrs



Fit formal childcare

Figure 4.A.5: Age youngest child 0–3 yrs

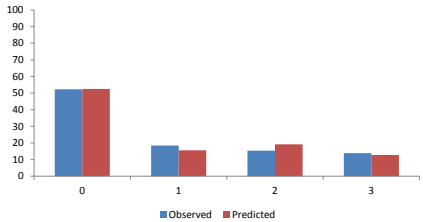
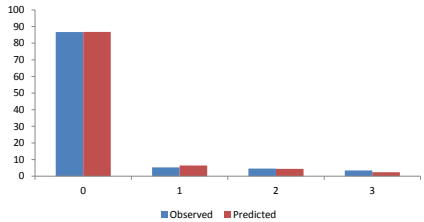


Figure 4.A.6: Age youngest child 4–11 yrs



5

LABOUR SUPPLY, FERTILITY AND CHILD CARE DECISIONS - A STRUCTURAL ANALYSIS OF FISCAL STIMULI FOR WORKING MOTHERS

This chapter is based on my job market paper.

5.1 INTRODUCTION

This chapter focuses on two long-standing challenges faced by many developed countries. The first is to increase labor participation of women, which is advocated as an important step towards narrowing the gender wage gap (Blau, 2012). The second is to promote fertility of the population, and thereby counteract the steadily falling birthrates. To meet these challenges, the governments are introducing public policies specifically tailored to increase female labor participation (in-work tax credits, child care subsidies, *etc.*), and others focused on raising fertility rates (universal child benefits, extension of maternity leave). However, as discussed by Apps and Rees (2009), childbearing has far-reaching consequences for mother's labor market attainment and *vice versa*, which means that a policy tailored to one of the policy goals is likely to have non-trivial spillover effects into the other domain of interest. These effects are generally unintended[1] and they can be often also counter-productive.[2] Some policies/fiscal stimuli are therefore bound to prove more (or less) desirable when investigated with both policy goals in mind.

In this paper, I develop a dynamic structural model of female labor supply, fertility and child care demand for households of married and cohabiting spouses. The model is intended for comparative analysis of labor-enhancing policies in light of the dichotomous policy goals presented above. The partenred women are assumed to be faced with the interdependent choices of working and having children. Furthermore, the spouses who decide to have children must also

1 A prominent example is the Earned Income Tax Credit (EITC) in the United States, which is a means-tested transfer for working families with children. Baughman and Dickert-Conlin (2003) find that the introduction of EITC indeed increased fertility levels, acknowledging that this was an unintended consequence of the labor-stimulating policy. Conversely, Eissa and Hoynes (2004) show that the program actually reduced labor participation among secondary earners in the household, since their participation was not required for eligibility and their earnings lowered the effective rate of the tax credit.

2 Baker, Gruber, and Milligan (2008) find that apart from increasing fertility, the extension of maternity leave had a negative effect on maternal labor supply. Lalive and Zweimuller confirm the adverse labor supply effects in the short-run, but they do not find evidence for their long-run persistence.

consider whether to use formal child care or not. The inclusion of child care choice in the household decision making is motivated by the apparent link between woman's market wotk engagement and her attitudes towards formal child care (See Apps, Kabátek, Rees, and Van Soest 2012), which can prove particularly important for the evaluation of public policies such as child care subsidies.

The evaluation of the effects of public policies on women's labor participation and fertility hinges on the assumption that these choices are dependent on financial incentives. Fortunately, the literature provides ample empirical evidence that these decisions are amenable to public policy interventions. The parental employment decisions have been shown to be responsive to a multitude of policy reforms, many of them being either in-work tax credits (IWTC) or child care subsidy reforms. The debate over relative efficiency of these fiscal stimuli is however by no means settled. The effects of child care subsidies on labor supply vary widely[3], whereas studies on in-work tax credits typically find sizable positive employment effects.[4] The fertility patterns are also generally found to be responsive to the financial incentives. Milligan (2005) shows that introduction of universal child benefits in Quebec had a large impact on fertility, leading to a 25% increase in fertility among women targeted for the full benefit. Similarly, Brewer, Ratcliffe, and dSmith (2012) show positive response of fertility to an increase of WFTC in the UK. Del Bono, Weber, and Winter-Ebmer (2012) show that job displacement and the subsequent loss of income stream significantly reduces fertility. Blau and Robbins (1989) find a negative link between fertility and child care costs faced by the US families. The child care demand is found to be responsive as well, increasing with more generous child care subsidies and higher incomes (Apps, Kabátek, Rees, and Van Soest 2012 and de Boer, Jongen, and Kabátek 2014).

In terms of dynamic structural models, there are two strands of literature with focus similar to this paper: models that explore female labor supply and fertility, and those that explore female labor supply and child care demand. Francesconi (2002) estimates a dynamic model of female labour supply and fertility. He shows that self-selection into fertility plays an important role for female labor supply, as he identifies negative links between maternal work ability and preference for children. Similar issues are investigated by Haan and Wrohlich (2011) who evaluate labor supply and fertility decisions of German women. In their paper, the dependency between maternal work and fertility is captured by allowing lagged choices to enter the

3 E.g. Lundin, Mörk, and Öckert (2008), Fitzpatrick (2010) and Havnes and Mogstad (2011) find no effect, Bettendorf, Jongen, and Muller (2012) find small effects, and Berlinski and Galiani (2007) and Lefebvre and Merrigan (2008) find large effects.

4 For an overview of studies analyzing the impact of the EITC in the US see Hotz and Scholz (2003), and for an overview of studies into the impact of the Working Families Tax Credit (WFTC) in the UK see Brewer and Browne (2006).

contemporaneous utility function. The second strand of literature is represented by Bernal (2008) and Griffen (2014) who develop dynamic models of female labor supply and child care demand. Apart from analyzing the substitution patterns of maternal and non-maternal care, they also explore what are the effects of different types of child care on the children's cognitive traits and later life outcomes. From the modeling perspective, it is interesting to note that both strands of literature are using bivariate choice models. The underlying emphasis is that the joint nature of work and fertility (or work and child care) is fundamental for sound policy analysis, whereas the third (omitted) choice variable is of secondary importance. However, since in both cases, the authors find strong ties between the choices considered, it seems to be appropriate to extended their analyses to a tri-variate setting, accounting for labor supply, fertility and child care in one joint framework.

The model presented here is estimated using a unique Dutch panel dataset which contains administrative records on earnings, child care expenditures and other employment-related information for a large sample of Dutch households. The records observed span the years 2001 to 2009, covering large reforms of family tax policy targeted at working parents. The reforms started in 2004 and within five years they gradually raised public spending on child care subsidies, and in-work tax credits (IWTC) for working parents. Such policy variation is convenient for the analysis since the family tax policy reforms (being strictly targeted at families with young children) changed the incentives underlying fertility, female labor participation and use of formal child care. As shown by Bettendorf, Jongen, and Muller (2012), the households did respond to the changing incentives, increasing the maternal labor supply and formal child care utilization. In my analysis I confirm these results and I show that the fertility was also stimulated by the reforms.

The policy analysis focuses on the fiscal stimuli which were introduced in the Netherlands within the observed period - the increase of in-work tax credits and child care subsidies. The results of the analysis show that the two policy reforms induce similarly strong changes of maternal labor supply as well as short-run government revenue. This means that according to the traditional labor supply metric, the policy makers could choose either policy without compromising budgetary efficiency. On the other hand, the response of fertility rates is less uniform. The fertility effects are more pronounced under the child care subsidy reform. However, the increased fertility is shown to raise the long-run costs of the child care reform, as more children become eligible for subsidized child care in the years after the implementation. The long-run maintenance costs of the child care subsidy reform therefore prove to be considerably higher than the immediate costs at the point of implementation. They are also higher than the long-run costs

of the in-work tax credit reform, which are shown to fall compared to the immediate costs.

This finding should be of particular interest for the policy makers, since the full costs of analyzed reforms will not be realized until several years after the implementation. This extra burden on the government budget is therefore likely to be overlooked by analyses which focus only on the short-run changes in the variables of interest, leading to misdirected policy recommendations.

The rest of the paper is organized as follows. Section 5.2 presents the structural model and the estimation strategy. Section 5.3 documents the institutional background and the two policy reforms which are exploited by the model. Section 5.4 describes the data and addresses some modeling choices which were adopted in Section 5.2. Section 5.5 presents the model fit, the estimation results and the outcomes of counterfactual simulations. Section 5.6 concludes.

5.2 STRUCTURAL MODEL

5.2.1 *Setup of the model*

The model evaluates households in a unitary framework, focusing on the decisions related to women and children. Husband's work allocation and disposable income are taken as given.[5] I further abstract from modeling formation of the households (marriage or cohabitation decisions), taking it as exogenous in the context of the model.[6] I do allow for accumulation and depreciation of mothers' human capital, which is argued to play a crucial role in mother's decisions to return into the labor force after taking maternity leave (Apps and Rees, 2009).

Timing of events starts with formation of the household (*i.e.*, the year when the spouses start living together) and ends when the woman reaches the retirement age of 65. The household starts making decisions as of period t=1. In each period, the woman is assumed to choose

5 This assumption is maintained throughout the literature (Francesconi 2002, Bernal 2008, Griffen 2014) and it is also supported by findings of de Boer, Jongen, and Kabátek (2014), who show that the labor supply elasticity of Dutch men with respect to changes in their spouses' earnings is rather small. The implications of modeling men's choices as a flexible part of household decision making are further discussed in Chapter 3, page 83.

6 Modeling of marriage/cohabitation markets is an interesting extension of the current setup of the model. Modeling such decisions would allow me to include single adults in the sample, which would broaden the scope of the analysis. Furthermore, assuming that marriage/cohabitation decisions are non-random and that they also react to financial incentives (such as tax credits for working parents), the reform-induced changes in the pool of married and cohabiting couples could influence the results of the current analysis. However, since I do not observe any pronounced differences in marriage rates throughout the reform period, I assume that the pool of households remains unchanged and I abstract from modeling these decisions.

her labor supply, which generates part of the disposable household income. At the same time the spouses decide whether to have children. This choice is restricted to couples where the woman is aged less than 45. A positive outcome of the fertility decision results in a childbirth in the next period. In the periods after the childbirth, the spouses also chooses whether to use formal child care or not. The outcome of this decision is applicable to all children in the household aged 11 years or less. At the point when all children become at least 12 years old and the woman also reaches the fertility threshold, the household decision problem simplifies to the market work decision of the woman.

The life-long fertility decisions are restricted by assuming that families can have at most two children.[7] In terms of the choice sets, I allow for three intensity levels of market work (full-time (FT), part-time (PT) and zero hours), four intensity levels of formal child care use (0 - 3 days per week), and a binary fertility decision. The choices corresponding to the formal child care do not contain 4 and 5 days of the service per week because these intensities are almost never observed.

The instantaneous utility is derived from household consumption c_t, woman's leisure time l_t, number of children aged less than 18 years N_t, fertility decision n_t (leading to a childbirth in the next period), and demand for non-maternal child care:

$$\begin{aligned} U_t = & \tfrac{{c_t}^{\mu}}{\mu} + \alpha_1^l l_t + \alpha_2^l l_t^2 + \alpha_3^l l_t N_t + l_t \varepsilon_t^l \\ & + \alpha_1^c fc_t + \alpha_2^c fc_t^2 + \alpha_3^c fc_t \cdot \mathbf{1}(age_{yng} < 4) \\ & + \alpha_4^c ic_t + \alpha_5^c ic_t^2 + \alpha_6^c ic_t \cdot \mathbf{1}(age_{yng} < 4) + \mathbf{1}\left(fc_t > 0\right) \varepsilon_t^c \\ & + \alpha_1^n N_t + \alpha_2^n N_t^2 + \alpha_3^n n_t + \alpha_4^n n_t l_t + n_t \varepsilon_t^n. \end{aligned} \tag{5.1}$$

The demand for non-maternal child care is expressed through two separate variables, distinguishing between the total hours of child care provided by formal carers fc_t (*i.e.,* kindergartens and daycare centers), and the total hours provided by informal carers ic_t (*i.e.,* grandparents and other relatives). This distinction is made to account for differences in costs and preferences for the two types of child care. Both variables are interacted with an indicator function which attains value one if the youngest child is less than 4 years old. This allows the model to account for differences in attitudes towards formal and informal child care between the households with school-aged children and the households with pre-schoolers. The stochastic part of the

7 This type of restrictions prevents families in our sample to be predicted to have excessive number of children. Selecting appropriate number of children at which the fertility should be capped is however non-trivial. An assumption which is often used in the literature is that families are allowed to have at most one child (see *e.g.,* Bernal 2008 or Brilli, 2013). Capping the total fertility at two children is a less stringent assumption, although it can be argued that it should be loosened further: 14.6% of Dutch families are observed to have three children whereas the total share of families with births of higher parities is not larger than 4%. This observation would suggest that three children fertility cap might be more appropriate.

utility function is represented by a triplet of choice-specific error terms $\{\varepsilon_t^n, \varepsilon_t^c, \varepsilon_t^l\}$ which accounts for choices that cannot be attributed to the observed factors.

The household's consumption c_t is assumed to be equivalent to household's disposable income y_t. I abstract from modeling the savings decision, so that the income has to be consumed in the same period it is earned[8]. The budget constraint takes the following form,

$$y_t = w_t \cdot h_t + y_t^h - T(w_t, h_t, \mathbf{X}_t) - p \cdot fc_t \cdot N_t^c + S(p, fc_t; \mathbf{X}_t). \quad (5.2)$$

Here, w_t is gross hourly wage, h_t is woman's labour supply and y_t^h are earnings of the spouse which are treated as given. $T(w_t, h_t, \mathbf{X}_t)$ denotes taxes and social security contributions. $p_t \cdot fc_t$ represents the cost of formal child care. If the spouses decide to use formal child care, it is assumed that the service is used by all children below 11 years of age, and therefore the cost is multiplied by the number of children in the appropriate age range, N_t^c.[9] $S(p, fc_t; \mathbf{X})$ denotes child care subsidies, which depend on total cost of formal child care and other household characteristics such as households' taxable income. In contrast, the costs of informal child care are not included in the budget constraint. The informal child care is often provided with no charge, or at a price that implies an unobserved subsidy from the carer. The lack of information about the costs of informal child care makes any effort to impute corresponding prices infeasible. The only way the costs of informal child care enter the model is through the parameters α_4^c, α_5^c and α_6^c in the utility function 5.1. The values of these parameters will represent an amalgamation of 1) intrinsic utility derived from this type of care, 2) effort required to find an informal carer, and 3) the unobserved costs of informal child care.

The woman's market wage is specified as follows

$$\log(w_t) = \beta_0 + \beta_1 educ + \beta_2 ex_t + \beta_3 ex_t^2 + \beta_4 educ_t \cdot ex_t + \varepsilon_t^w, \quad (5.3)$$

8 This assumption can be interpreted in the following way: Households are restricted in their inter-temporal consumption smoothing, since they are not able to save or borrow against their future income. In real life, such borrowing constraint is unlikely to be universally applicable, but it may well describe the situation faced by many families with young children. Allowing households to borrow and save would change the nature of the decision problem faced by the spouses. It would alleviate the severity of realized negative income shocks, enabling households to sustain prior levels of consumption and child care use (or to follow through with their family planning) even while facing adverse economic conditions. If the households in our sample are indeed actively smoothing their consumption, then allowing for borrowing and saving would strengthen the link between disposable income and the choices which are being modeled.

9 There is, however, an important distinction between the type of child care provided to children aged 0-3 (*daycare*) and to children aged 4-11 (*out-of-school care*). Since the latter group attends school for half of the day, they require half the child care provision needed for younger children. The daily formal child care costs for older children are therefore halved compared to the costs for pre-schoolers. Further discussion of the formal child care costs will be provided in the Section 5.4.

so that it explicitly accounts for heterogeneity among earnings profiles characteristic for specific education levels. It further allows for the effects of human capital through the dependence on cumulative experience gathered over the five years preceding the current period,

$$ex_t = \sum_{i=1}^{5} \mathbf{1}(h_{t-i} = 1). \tag{5.4}$$

That way, I account for appreciation/depreciation of labor skills in a span of years which is likely to be highly relevant for the wage determination process. The span of five years was chosen for convenience, since the individual labor market histories are only known for 5 years prior to the period when all the households' decisions are observed for the first time. An alternative approach would be to allow for lifetime experience accumulation (see Francesconi, 2002, or Bernal, 2008), where the experience entering the wage equation represents every labor market engagement preceding the given point in time. An issue with this approach is that the lifetime experience is rarely observed and so it has to be approximated by potential experience, which is equal to worker's age minus her years of schooling. This approximation may prove particularly problematic in the context of female labor supply, where women often opt out of the labor market to rear their children. The potential experience would therefore often overestimate their actual experience. For that reason, I prefer to rely on the former approach, drawing the inference based on the experience accumulated during the observable period.

The household is assumed to predict evolution of woman's future earnings according to the equation 5.3. Furthermore, in order to form expectations about the disposable household income, the spouses are also assumed to predict future realizations of man's earnings. This prediction is based on another Mincer-type equation, estimating the man's earnings from his experience and educational attainment,

$$\log(y_t^h) = \gamma_0 + \gamma_1 educ^h + \gamma_2 ex_t^h + \gamma_3 \left(ex_t^h\right)^2 + \gamma_4 ex_t^h \cdot educ^h + \varepsilon_t^h. \tag{5.5}$$

The woman's time allocation comprises of market work h and leisure l, and it is subject to the following constraint:

$$h + l = TC, \tag{5.6}$$

where TC is a time constraint amounting to the standard time endowment of 16 hours per day. The full-time and the part-time work allocations take up respectively 8 and 4 hours per day, which leaves at least eight hours per day for the leisure endowment. It should be noted that the leisure here is defined in very broad terms, accounting both for the actual leisure time, but also for the non-market work, and the maternal child care.

I also specify a second time constraint which requires the sum of all child care inputs to exceed minimal time requirement TR.

$$l + fc + ic \geq TR(age_{yng}). \tag{5.7}$$

The value of TR is dependent on the age of the youngest child, amounting to 16 hours per day for pre-schoolers and 12 hours per day for school-aged children younger than 11 years.[10] This constraint maintains that the child cannot be left unattended at any time during the day, avoiding situations where parents are predicted to work full-time without having anyone to look after their offspring. The minimal time requirement also facilitates the incorporation of informal child care into the structural model. Since the data used for this analysis come from administrative records, I do not observe any information on the provision of informal child care ic. Therefore, I assume that the informal care supplied to the given family is equal to the difference between the time requirement TR and the sum of formal child care and maternal leisure,

$$ic = \max\left(0, TR(age_{yng}) - fc - l\right). \tag{5.8}$$

This means that the inequality in equation 5.7 will be applicable only for those women whose sum of formal child care and leisure time alone is greater than TR. All the other cases are bound to satisfy the equation with equality.

The stochastic components $\varepsilon_t = (\varepsilon_t^l, \varepsilon_t^c, \varepsilon_t^n, \varepsilon_t^w, \varepsilon_t^h)$ are assumed to be drawn from a multivariate normal distribution with zero mean vector and diagonal covariance matrix Σ. The errors are drawn independently of their previous realizations.[11] The state space at time t is defined as a set of all the relevant factors which affect either current or future values of household's utility,

$$\Omega_t = \{ex_t, ex_t^h, educ, educ^h, y_h, p_t, age_{yng}, age_{old}, N_t, \varepsilon_t\} \tag{5.9}$$

Mother's decision problem can be then written in terms of value functions $V_t(\Omega_t)$, which maximize over expected present value of lifetime utility conditional on the realized state Ω_t,

$$\begin{aligned} V_t(\Omega_t) &= \max_{k \in K(t)} \left[V_t^k(\Omega_t)\right] \\ V_t^k(\Omega_t) &= U_t^k(\Omega_t) + \delta E\left(V_{t+1}(\Omega_{t+1}) | \Omega_t, d_{kt} = 1\right), \quad t < 65 \\ &= U_{65}^k(\Omega_{65}), \qquad\qquad\qquad\qquad\qquad\quad t = 65 \end{aligned}$$

10 Children older than 11 years are assumed not to require any child care.

11 The assumptions imposed on the error terms can be relaxed to allow for more flexible treatment of unobserved heterogeneity. Since we already have normally distributed errors, the diagonal terms of the covariance matrix Σ can be allowed to differ from zero. That way we would account for mutual dependence of the stochastic components at the same point of time. Another way is to allow for temporal dependence of errors, such as in Blundell, Dias, Meghir, and Shaw (2013) where the errors are assumed to follow an AR(1) process with i.i.d. innovations.

5.2.2 *Solution and estimation of the model*

The model is solved numerically, using backward recursion to solve the *Emax* functions $E\left(V_{t+1}(\Omega_{t+1})|\Omega_t, d_{kt}=1\right)$. The functions are evaluated for every choice k in the choice set K and for every element of Ω_t, starting from the last period T. The solution is conditional on given parameterization of the structural equations, so that the *Emax* functions serve as an input into the model optimization procedure.[12]

The value functions for each alternative are known conditional on the deterministic part of the state space. The stochastic vector ε_t remains unobserved. In order to quantify probability of choosing the observed choice, it is necessary to integrate over the joint distribution of ε_t, isolating the realizations of stochastic shocks which render the observed choice k to be the alternative with the highest value function. To avoid issues with optimization over non-smooth probability spaces, we employ the kernel smoothed frequency simulator proposed by McFadden (1989).

The probability of any sequence of choices made by the woman can be therefore computed as follows,

$$\Pr(k_{A_0}, ..., k_T|\bar{\Omega}_{A_0}) = \prod_{t=A_0}^{T} \Pr(k_t|\bar{\Omega}_t) = \prod_{t=A_0}^{T} \Pr(V_t^k > V_t^j, \forall j \neq k|\bar{\Omega}_t) \tag{5.10}$$

where the choice probabilities are explicitly conditioned on the observable part of the state space $\overline{\Omega}_t$.

5.3 INSTITUTIONAL SETTING

The Dutch family tax policy has been recently subject to two major reforms: a reform of child care subsidies and a reform of in-work tax credits (IWTC) for working parents. Both reforms were gradual and they were being implemented throughout the years 2004 to 2009. The result of these reforms was a substantial increase of public spending on families with small children. The reforms were aimed at increasing maternal labor force participation, whereas the notion of using these stimuli to increase fertility rates was not made explicit. The incentives

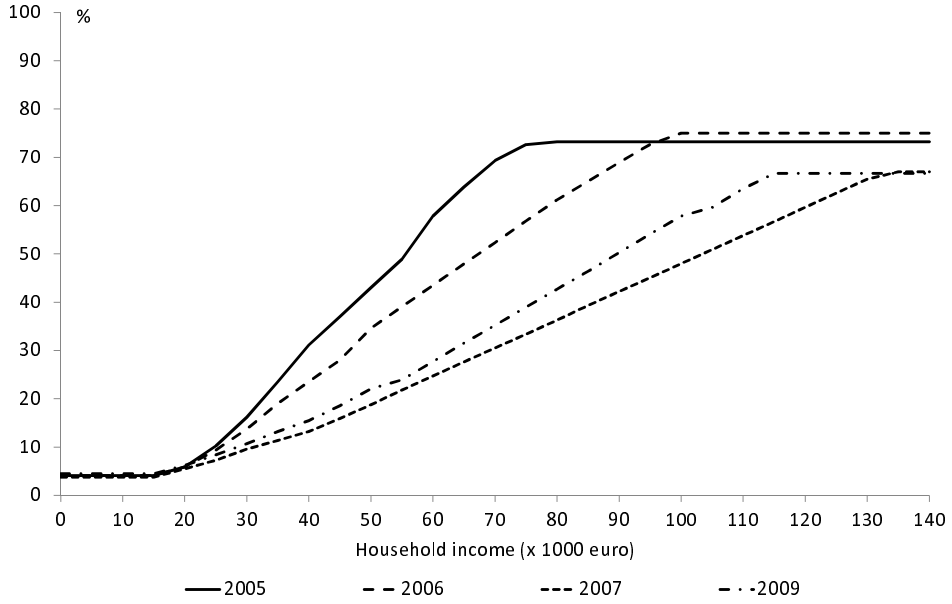
12 The state space considered is, however, very large. The decision makers have to be aware of the age structure of children, cumulative experience, and other factors which render solving the decision paths of entire state space infeasible for practical purposes. To overcome this, I employ estimation strategy developed by Keane and Wolpin (1994), the regression based interpolation. The emax functions are evaluated at a random subset of choice and state variable combinations and I interpolate the values of Emax functions for the rest of the state space. The interpolation step consists of running a series of regressions with the regressand being emax functions which are drawn for the given time period, and the regressors being the variables entering the observable part of the state space.

for childbearing however changed as well, since both the child care subsidies and the in-work tax credits were targeted at the parents who had at least one child aged 12 years or less.

5.3.1 *Child care subsidy reform*

The child care subsidy reform started with the introduction of Law on Child care (*Wet kinderopvang*) in 2005. Immediate effect of the law involved formal changes in the institutional setting of child care subsidies, eliminating regional differences in subsidy rates and other idiosyncrasies characteristic for the old system (see de Boer et al. 2014). As of 2005, all families who used formal child care (for their children aged 0 to 11) qualified for the same subsidy scheme which had been set by the central government.[13] The institutional changes in 2005 however changed little in terms of the effective subsidy rates faced by the majority of Dutch families (see Plantenga, Wever, Rijkers, and de Haan 2005 and Ministry of Finance 2010).

Figure 5.1: Parental contribution rate to the child care costs for the first child



Source: own calculations using publicly available subsidy tables.

More important were the changes that followed in 2006 and 2007. During this period the child care subsidy rates got substantially higher, with the largest increase taking place in 2007. Figure 5.1 shows the changes in the parental contribution rate for the 'first child'.[14] First, note that the parental fee depends on the income of the household. In

13 The unification of subsidies also allowed for collection of the data on child care use which are used in this analysis.

14 The Tax Office defines the first child as the child for which the parents have the highest child care expenditures.

all years, households with the lowest income received the highest subsidy (up to 96% of the full price) and this subsidy rate hardly changed throughout the reform period. For the middle income households the subsidy rate went up by 20 to 40%-points, whereas the increase in the subsidy for the highest income households was somewhat smaller. On average, the child care cost share paid by the parents dropped from 37% in 2005 to 18% in 2007 (de Boer et al. 2014). Next to the drop in parental fees, from 2007 onwards schools were obliged to act as an intermediary for parents and child care institutions to arrange out-of-school care. In 2008 there were virtually no changes in child care subsidies, but then 2009 witnessed a partial reversal of the policy change, and parental fees were raised somewhat compared to their prior levels. To illustrate that the parents were responding to the incentives, Table 5.1 shows descriptive statistics of the use of child care in the period 2006-2009.

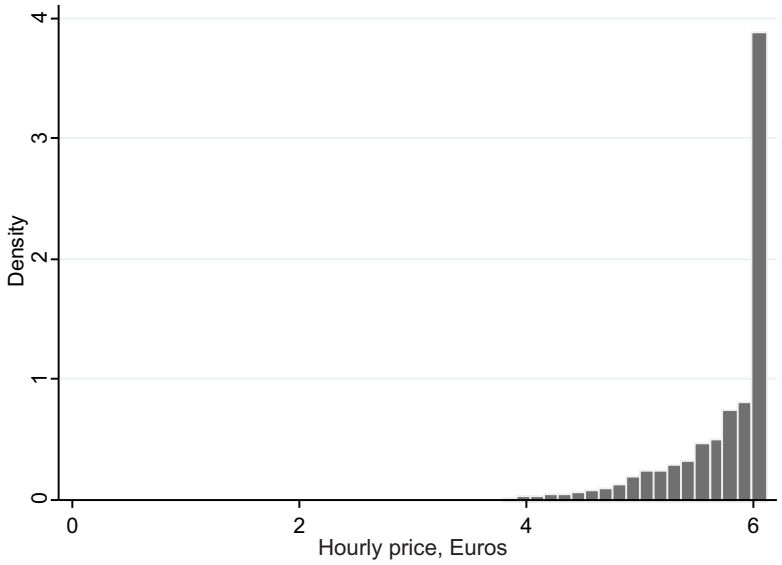
Table 5.1: Use of child care 2006-2009

	Daycare			
	2006	*2007*	*2008*	*2009*
Share households	29.8	40.6	49.8	53.8
Average number of children	1.3	1.3	1.3	1.3
	(0.5)	(0.5)	(0.5)	(0.5)
Average number of hours per child	19.5	18.7	19.5	20.0
	(9.3)	(10.4)	(10.3)	(10.4)
	Out-of-school care			
	2006	*2007*	*2008*	*2009*
Share households	7.1	11.3	15.2	18.4
Average number of children	1.4	1.4	1.5	1.5
	(0.6)	(0.6)	(0.6)	(0.6)
Average number of hours per child	9.6	8.7	9.4	9.9
	(6.4)	(6.3)	(6.8)	(7.6)

It is also important to discuss the situation on the supply side of the formal child care provision, which motivates several modeling choices adopted in the empirical analysis. As documented by Akgunduz and Plantenga (2013), the number of child care centers in the Netherlands, as well as their size, expanded rapidly in response to the reform. This attenuated problems with rationing (which has been observed in other European countries, see Kornstad and Thoresen 2007 or Haan and Wrohlich 2011). Akgunduz and Plantenga (2013) show that despite its swiftness this expansion did not impede the quality of provided care. It is also worth noting that the real prices of formal child care remained stable throughout the reform period. This goes against the general

equilibrium effects, which suggest that the price should increase since the subsidy induces more demand for formal child care. The apparent price stability might have been achieved by putting a cap on the maximal hourly price of child care which qualifies for the government subsidies. This cap was initially set to 6.03 Euro, and a large majority of the child care centers decided to set their hourly price equal to the threshold (See Figure 5.2). This standardization of prices probably contributed to the stability of child care quality indicators observed by Akgunduz and Plantenga (2013). In the following years, the price cap has been increased somewhat to match the Dutch consumer price index.

Figure 5.2: Histogram of hourly prices of formal child care in the Netherlands, 2006-2009, in Euros



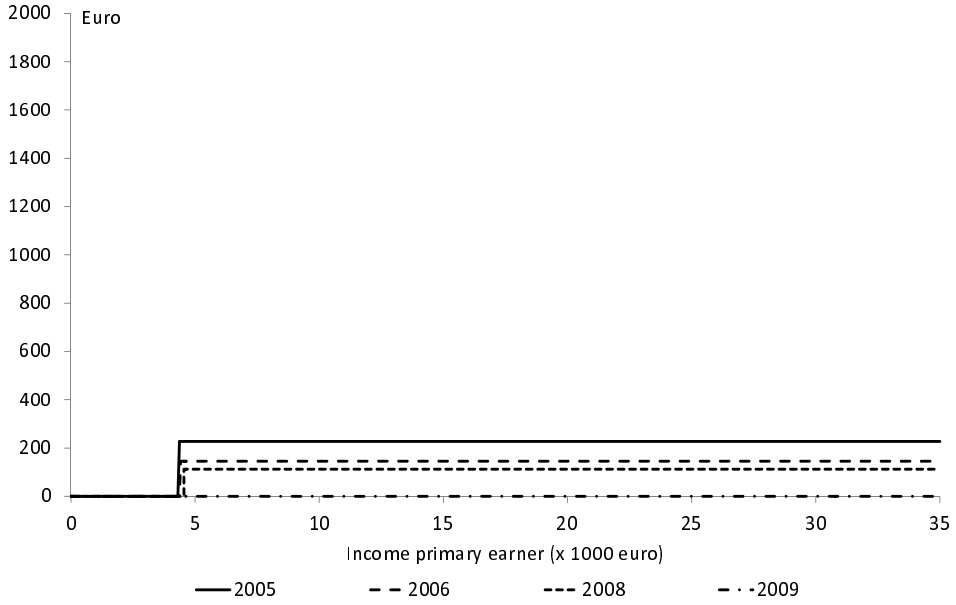
5.3.2 *In-work tax credits reform*

The reform of in-work tax credits for working parents started one year earlier than the child care subsidy reform. Under the original system, a uniform tax credit of 220 Euro was provided to every working parent who earned more than 4366 Euro *p.a.* and whose youngest child was up to 12 years old. Throughout the years 2005-2009 this credit was phased out (See Figure 5.3), and it has been gradually replaced by a more targeted IWTC policy which was introduced in 2004. This new tax credit remained conditional on the child's age, but it was no longer awarded to both spouses. It was provided only to the secondary earner in the household, defined as the spouse with the lower labor income.[15] Effectively, this meant that the tax credits

15 Single parents were treated as secondary earners.

became targeted predominantly at women, and the men's part was gradually retracted. Figure 5.4 shows the changes in the amount of total tax credit applicable to the secondary earners in years 2004-2009. The tax credit started as a flat-rate subsidy amounting to 514 Euro. This rate gradually increased to 858 Euro in 2008, which was followed by a more profound change of the policy in 2009. As of 2009, the amount of tax credit awarded to the secondary earner has been made dependent on her labor earnings. The credit increases with earned income up to a maximum amount which is not phased out at higher incomes. This adjustment was intended to stimulate full-time work arrangements among mothers with children (who were often observed to engage in small part-time jobs) In 2009, the maximum credit of 1,765 Euro was reached at 30,803 Euro of gross individual income (the minimum wage of a fulltime worker during the same period was 16,776 Euro).

Figure 5.3: EITC for primary earners with children

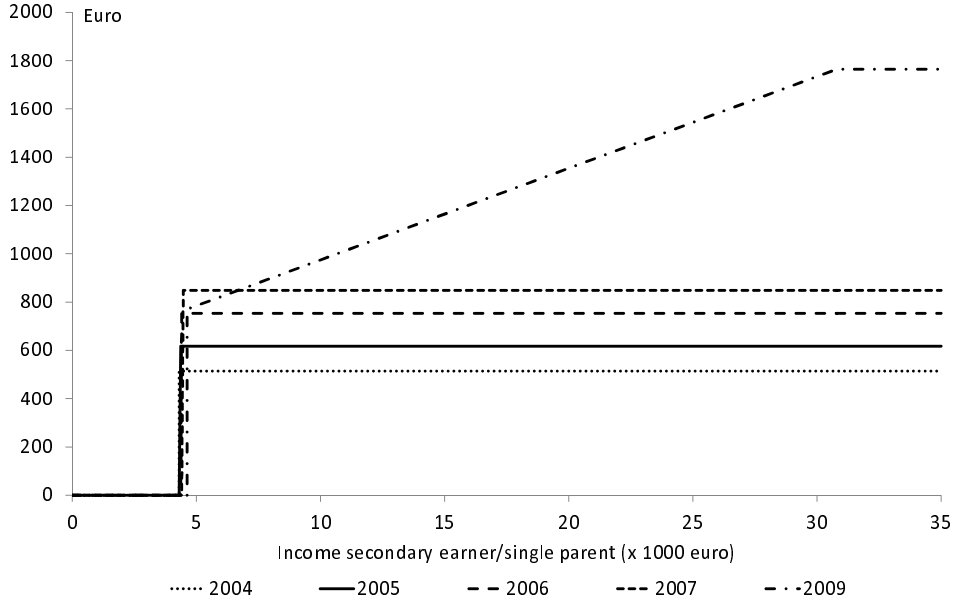


Source: Tax Office.

5.3.3 *Modeling of the reforms*

The policy changes described above show that the Dutch families have been subject to many changes in the structure of incentives underlying employment, child care demand and fertility. This variation facilitates identification of the model, since it introduces sizable and arguably exogenous variation in the incentives underlying the decisions considered in our model. The costs of maternal leisure changed in several ways, being affected directly through the targeted IWTCs, and indirectly through the child care subsidies and the loss of father's tax

Figure 5.4: EITC for secondary earners with children and single parents



Source: Tax Office.

credits. Also the costs of formal child care and the costs of childbearing fell considerably due to the reforms.

In the context of the model, it is assumed that the households are unaware of the prospective policy changes, assuming that the current policy regime will be maintained indefinitely. During the first reform year, they will update their expectations, replacing the prior policy regime by its current variant. This belief-updating process continues until they reach the final reform year, after which the policies are assumed to be kept unchanged.

5.4 DATA

The model is estimated using a unique panel dataset provided by Statistics Netherlands (*CBS*). The dataset is built up from several administrative sources, combining information from the Labour Force Survey, Social Statistical Panel, administrative data from municipalities, and Formal Child care Database of the Tax Office. The resulting panel constitutes a comprehensive source of information over a large sample of Dutch population (1.05 million individuals) who were observed in the period 2001-2009. The dataset contains detailed information on socio-economic characteristics of the individuals, household structure, labor market outcomes, and child care utilization for families with young children. The child care information is however unvailable for years prior to 2006, since there has been no standardized child care subsidy in place (and the tax office therefore did not collect household-level child care data).

In terms of sample selection, several rules have been applied to extract the sample used for the analysis: Due to the limited availability of child care information, I restrict myself to years for which all choices are fully observable (2006-2009). The information from prior periods is however not discarded - it is used to determine correct experience levels in the equation 5.4. The sample is restricted to complete families only. This restriction follows from the fact that the decision making problem of single parents is likely to be very different from the one of complete families (Blundell and Shephard, 2012). Furthermore, the data on single parents often lacks important sources of non-labor income, such as funds from grandparents, or alimonies. Similarly to Chapter 4, I exclude cases where where at least one of the spouses is self-employed, student, disabled, or on unemployment benefits. This is done since I cannot determine either the budget constraint or the choice set corresponding to these household types. I also exclude the families where I observe only one of the spouses, and families which separated within the observed period.

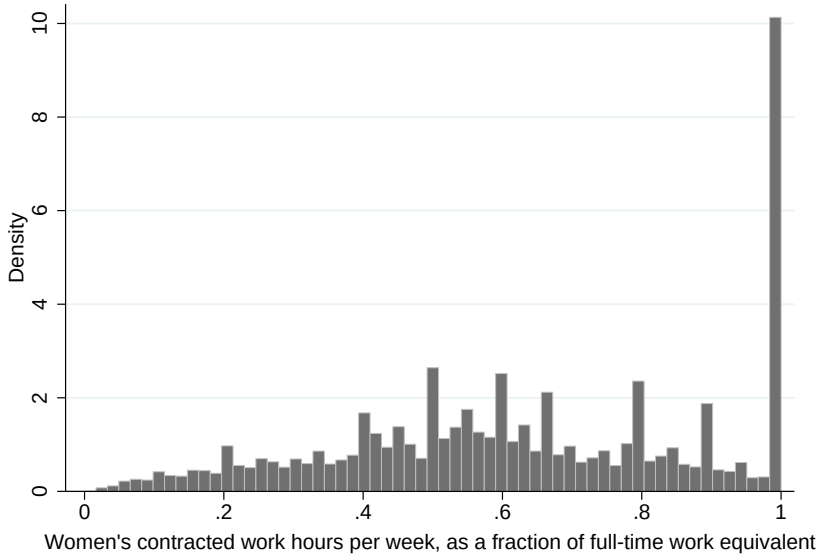
Finally, the dynamic model is unlikely to require the full scope of the administrative dataset in order to provide stable estimates of preference parameters. In order to make the estimation computationally feasible, I restrict the sample to a random subset of the data that remains after applying the selections listed above. The subset which is used for the presented empirical analysis contains 5000 households. This sample has been deemed sufficiently large to ensure stability of both the estimated structural parameters and the counterfactual simulations.[16]

A set of descriptive statistics corresponding to the final sample is provided in Table 5.2. The personal characteristics are presented for both genders separately, illustrating the relative roles and traits of the spouses in the analyzed families. These records highlight some important idiosyncrasies of the Dutch data. Firstly, the female labor participation rate of 0.78 is relatively high. In the international context it is slightly below the rates of Nordic countries which are recognized for active engagement of the female workforce, and slightly above the rates of English-speaking countries.[17] The average hours worked show that women, if employed, are working predominantly in part-time arrangements. The Dutch labor market is considered to be highly flexible in accommodating the needs of part-time workers, which is documented by the diversity of observed work hours shown in the Figure 5.5.

16 A robustness check based on a newly drawn sample of 10000 women was performed and the estimated structural parameters were not significantly different from those presented in Section 5.5.2.

17 In 2006, the aggregate female participation rate was 0.68 in the Netherlands, 0.72 in Norway and Denmark, and 0.66 in the US and the UK. The EU-wide average was 0.57 (OECD, 2014)

Figure 5.5: Histogram of contracted work hours per week, expressed as shares of the full-time work equivalent, women aged 18-65



The Figure plots the histogram of contracted work hours expressed as a fraction of full-time work equivalent (fte) which generally amounts to 40 work hours per week. The spike representing full-time (1 fte) workers accounts for 16% of female workforce.[18] The fact that women are working mostly part-time is also reflected in their annual earnings, which are 40% lower compared to the earnings of their partners.

Table 5.2 further shows how many families are actively using formal child care. As noted in the Section 5.2, a distinction is made between daycare (children 0–3 years of age[19]) and out-of-school care (children 4–11 years of age).[20] According to the statistics, the daycare services are used much more than the out-of-school care. This difference in attitudes towards child care motivates the inclusion of age-dependent child care parameters in the utility function (equation 5.1). Indeed, many modeling assumptions about child care are closely related to the idiosyncrasies observed in the data.

The Tax Office reports a very precise breakdown of annual hours of formal child care used, separating different types of formal child care used by each child in the household. Here, these quantities were re-

18 For the analysis, I use a broader definition of full-time work. Full-time workers are defined as those women who work 0.8–1.0 fte. Part-time work is equivalent to 0.2–0.799 fte and the remaining women are coded as working 0 hours.

19 Maternity leave in the Netherlands is rather short, 3 months after the birth of the child, which can be supplemented with 3 months of parental leave for which the replacement rate is however rather low (OECD, 2007).

20 Children in the Netherlands go to school when they turn 4, and are typically in secondary school when they are 12 years of age.

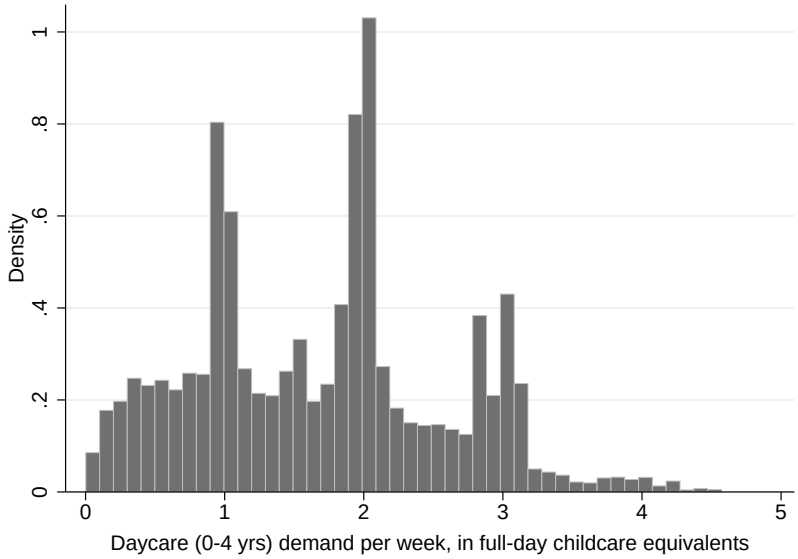
Table 5.2: Descriptive statistics stratified by gender

	Men	Women
Age	40.4	38.0
Labor participation rate	0.96	0.78
Hourly wage	21.3	16.2
Hours worked per week[a]	38.9	22.1
Native	0.84	0.83
Western immigrant	0.08	0.09
Non-western immigrant	0.08	0.08
Lower educated[b]	0.21	0.18
Middle educated[b]	0.42	0.50
Higher educated[b]	0.37	0.32
Children between 4 and 12 years	0.54	0.54
Urban area (share)[c]	0.17	0.17
Non-urban area (share)[c]	0.83	0.83
Daycare use (share)[d]	0.31	0.31
Out-of school care use (share)[e]	0.13	0.13
Number of observations	913236	913236

[a]Hours worked per week by the employed. [b]Education is classified as follows (using the Dutch abbreviations): i) lower educated = BO and VMBO, ii) middle educated = MBO, HAVO and VWO, iii) higher education = HBO and WO. [c]An urban is defined as large (small) when it has 150,000 inhabitants or more. [d]Daycare use among families with at least one child aged 0–3. [e]Out-of school care use among families with at least one child aged 4–12 years.

scaled to represent weekly demands for daycare and out-of school care for every child in the respective services. Figure 5.6 shows that the vast majority of families are observed to use less than 4 full days of daycare per week. Furthermore, the spikes of the distribution illustrate that it is commonplace to use full-days of daycare (rather than half-days or other quantities).

Figure 5.6: Histogram of household's demand for daycare per week, expressed in full-day child care equivalents per week, active users only

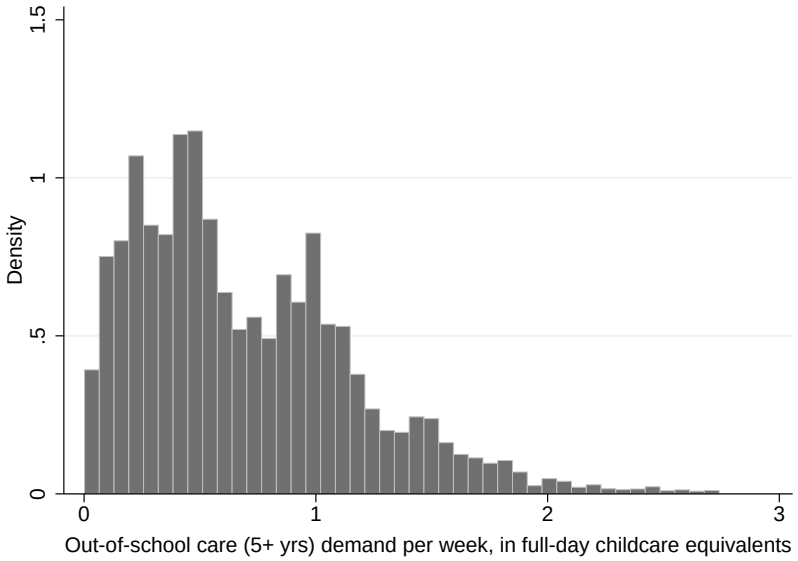


Similar picture emerges also for the out-of-school care demand in Figure 5.7. The reported hours are around 50% lower compared to daycare, which follows from the fact that school-aged children do not need more than half-days of child care. Also in this case, the histogram indicates that families tend to use out-of-school care for less than four days per week (which is equivalent to two full-days of child care), and the demand bunches around rounded half-days of care. These findings lend credibility to the modeling assumption that women are choosing among four formal child care alternatives, corresponding to 0–3 days of formal child care per week. The data also support the hypothesis that the child care choice applies equally to all children in the household. In 91% of the cases, families use overlapping child care arrangemets for each of their children aged less than 12.

In the context of the model, the observed demands for formal child care are translated into corresponding discrete choices by taking weekly household averages[21] and rounding them to the nearest integer.

21 The formal child care demands are averaged over all children in the household who are less than 12 years old. The child care demand of school-aged children is multiplied by two, so that the resulting average could be interpreted as a number of

Figure 5.7: Histogram of household's demand for out-of-school care, expressed in full-day child care equivalents per week, active users only



This rounded value represents the number of days of child care the family has decided to use.

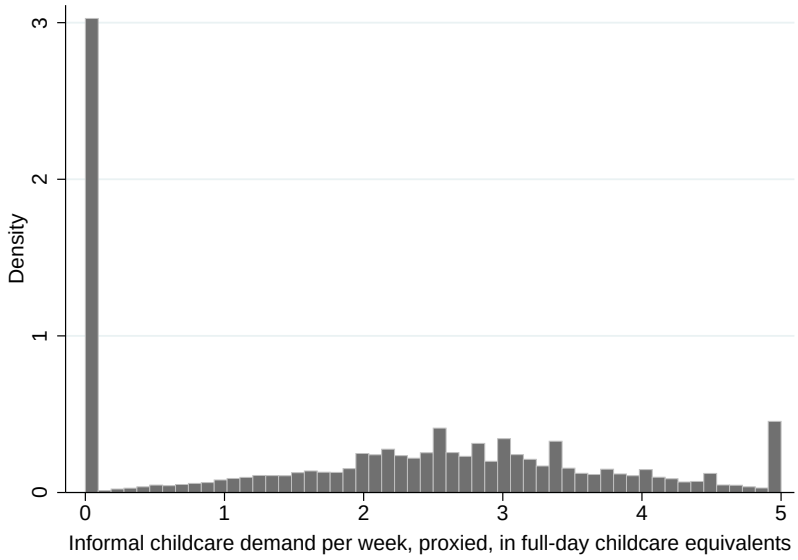
The observed patterns of formal child care highlight another important fact about Dutch families. That is, the households are rarely fully reliant on the services of child care centers and kindergartens. The informal child care is an important source of child care provision in the Netherlands.[22] The informal providers are often family relatives, friends, neighbors, or husbands themselves (although men in the sample are rarely observed to change their work arrangements after a child is born). In order to get more concrete grasp on the extent of informal child care use, I employ the equation 5.8 and derive the informal child care proxy ic from the observed intensities of maternal work and formal child care provision. Figure 5.8 presents the historam of proxied informal child care. The spike at zero represents 26.7% of households which were assigned no informal child care at all. This means that the observed combination of formal child care provision and mother's leisure time covers the entire time required by children in these households. The histogram of non-zero hours peaks firstly around 2.5 days and then again at 5 days, accounting for those women who work full-time and do not use formal child care. For modeling purposes, the informal child care provision is treated in the following way: Instead of using the actual quantities of formal child care and

days during which the children were taken care of by formal child care providers, instead of mother or other carers.

22 According to the 2008 wave of Dutch LISS socio-economic panel, 65% of families with pre-school children were regularly using some form of informal child care.

market work hours to derive the informal child care provision, I use the values corresponding to the discrete choices which represent the observed quantities. This step is necessary in order to assign standardized values of informal care to all alternatives that can be chosen by the households (since we observe the actual quantity of informal child care for one of them only).

Figure 5.8: Histogram of household's demand for informal child care, expressed in full-day child care equivalents per week



5.5 RESULTS

5.5.1 *Reduced form analysis*

The policy reforms of 2004-2009 were associated with profound changes of incentives underlying work, child care and fertility decisions of the Dutch households. Under the assumption that families do respond to such incentives, it should be possible to get some insight about the reform effects by conducting a reduced form analysis of the households' choices prior to and within the reform period.

Table 5.3 presents estimated parameters of four household-level Random-Effects models, each focusing on one of the household's decisions which could be influenced by the manifested reforms. The dependent variables are binary indicators representing either the incidence of specific events or choices within the given year. These are: 1) realized marriages among single women 2) childbirths among the child-less couples, 3) use of formal child care among the parents with children younger than 12 years, and 4) labor participation of partnered women aged 18 - 65. The independent variables include a

set of socio-economic characteristics, education level dummies and yearly dummies.[23]

Table 5.3: Dynamics of the main variables of interest over the period 2001-2009, Random-Effects Regression

	(1) Marriage	(2) Fertility (1st time parents)	(3) Child care Use	(4) Market Work
Husband's earnings (log)		0.0232***	0.0153***	0.00164***
PT work	0.00369***	-0.0460***	0.221***	
FT work	0.0198***	-0.397***	0.0126***	
Age	-0.00581***	0.193***	0.143***	-0.0425***
Age^2	0.00003***	-0.00299***	-0.00215***	0.000533***
Immigrant	-0.00491***	0.00836***	-0.0172***	-0.0727***
Grandparents in the h'hold	0.00371***	0.566***	0.166***	-0.0510***
No. of adults in the h'hold	-0.00333***	-0.00871***	-0.00118	-0.00629***
Year: 2002	0.00137***	0.00239		0.00713***
Year: 2003	0.00218***	0.00121		0.00770***
Year: 2004	0.00119**	0.00910***		0.00573***
Year: 2005	-0.000159	0.00578**		0.00760***
Year: 2006	-0.000231	0.0117***		0.0170***
Year: 2007	-0.00144**	0.0237***	0.0776***	0.0308***
Year: 2008	-0.00102	0.0290***	0.130***	0.0381***
Year: 2009	-0.000171	0.0366***	0.158***	0.0390***
Education dummies	yes	yes	yes	yes
Constant	-0.267***	-2.749***	-2.591***	1.787***
Observations	534768	334380	117112	913236

*** p<0.01, ** p<0.05, * p<0.1

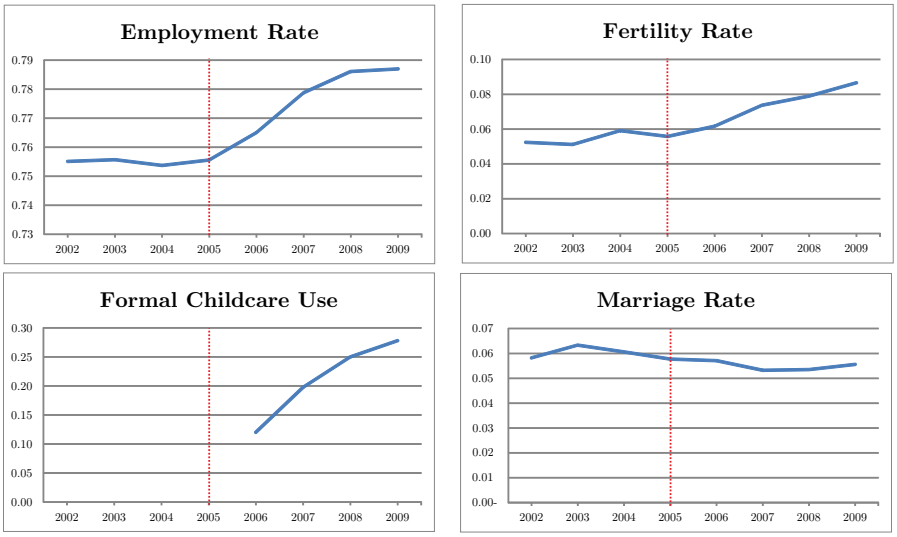
The covariates included in the regression specifications exhibit plausible signs, and given the number of observations they are rarely insignificant. The yearly dummies illustrate how did the analyzed decisions change prior to and within the reform period. The results in the first column show that the propensity to get married remained rather stable throughout the whole span of eight years. We observe a slight peak around 2003 and then a gradual decrease within the main reform period 2005-2009. The second column corresponds to the incidence of births among childless couples. This subsample was chosen for the sake of exposition, because the patterns observed among the first-time parents illustrate the changes in childbearing without the confounding effects of birth spacing and other issues related to higher-order births. Within the span of our data, we can see steady rise of childbirth incidence among childless, which mimics the gradual introduction of child care subsidies and in-work tax credits. The spike

23 It should be noted that since this analysis is purely reduced-form, the size of the estimated effects may be confounded by interdependencies of the considered choice variables, and also by their relation to other factors which are not accounted for here (*e.g.*, business cycle or local unemployment rates). Nevertheless, it can still give us valuable insights about general direction of the reform effects.

in 2004 can be potentially attributed to the preceding peak of marriage rates in 2003.

Time dummy coefficients in the third column lead to the same conclusion as the statistics presented in Table 5.2 - raising child care subsidies spurred the demand for formal child care. Since I observe the child care demand only for years 2006-2009, the regression is restricted to this period, with year 2006 being the baseline. The last column documents that the women also increased their labor supply attainment throughout the reform period. According to the results, the female labor participation began to grow in 2005 and by the end of the reforms it was almost four percentage points higher, compared to the 2001 levels. Similar picture emerges if we plot the results graphically. Figure 5.9 shows the development of the analyzed variables throughout the span of the data.

Figure 5.9: Dynamics of the main variables of interest over the period 2002-2009, plots of initial levels and subsequent changes



This preliminary analysis, however descriptive, leads to several important findings. Within the reform period, the households do exhibit changes in their allocations which are endogenous to the structural model. The changes are also in line with the economic theory (rising female labour participation, child care demand and fertility). Furthermore, the households do not exhibit changes in the marriage decisions which are accordingly taken as given by the model. Such outcomes are supportive of the assumptions underlying the structural model and the results presented in the next section.

5.5.2 *Structural analysis*

The model is parameterized according to the equations presented in the Section 5.2. An overview of estimated parameters is given in Table 5.4. The signs of parameter estimates are consistent with individual optimization behavior. The linear coefficients are positive, implying that mothers do derive utility from initial consumption, leisure, child care, and fertility. The negative signs of quadratic coefficients show that the utility gains for these variables exhibit decreasing returns in scale. The parameter for interaction of formal child care and leisure is negative, reflecting the fact that formal child care is usually used as a substitute for maternal time with the children.

Table 5.5 presents an assessment of goodness of fit of the model. I use the estimates of structural parameters to simulate choice paths and earnings trajectories for the given sample of women, starting from the observed initial conditions.

The simulated paths and earnings are used to derive a set of sample moments corresponding to the artificial data. The simulated moments are compared with their observed counterparts. The results in Table 5.5 show that the simulated and observed moments are matched very well. To account for the fact that the variables of interest evolve over time, Figure 5.10 shows their observed and simulated averages for all years considered. The plots show that the model prediction closely matches the observed patterns. This means that the model is capable of replicating the effects induced by the two policy reforms, which can serve as a validation of the estimated structural parameters.

Another potential concern is the behavior of predicted earnings over time. It is often found in the literature that the dispersion of earnings increases as the workers grow older (Haan and Wrohlich, 2011), which is considered as an argument for modeling wages as a Markov process. I argue that substantial part of this dispersion can be captured by the specification of the two earnings functions, 5.3 and 5.5. The dependence of predicted wage on recent work experience allows for higher earnings in the later periods, but it also allows for larger variance of earnings since some people will accumulate less experience than others. Figure 5.11 shows gender-specific plots of wage dispersion against worker's age. The two trajectories present in the plots correspond to the dispersion observed in the data and the dispersion predicted by the model. The gap between observed and predicted dispersion is attributed to the error terms, since the observables can predict only a part of the total wage variance. Adding the estimated variance of the wage errors to the predicted dispersion would bring the two lines very close together. Both trajectories are shown to follow similar patterns until the age of 50, when the observed dispersion starts to diverge. This divergence is most likely driven by

Table 5.4: Parameterization and Parameter Estimates

	Value	St.dev.
utility function		
CRRA parameter (μ)	0.7146	(0.0311)
leisure (linear)	1.771	(0.0237)
leisure (quadratic)	-0.1634	(0.0509)
formal child care (linear)	1.9411	(0.1479)
formal child care (quadratic)	-0.4128	(0.1542)
formal child care $*$ preschool	0.1301	(0.0671)
formal child care $*$ leisure	-0.0622	(0.0092)
informal child care (linear)	1.0832	(0.6276)
informal child care (quadratic)	-0.0127	(0.0043)
number of children	0.5622	(0.0092)
number of children (quadratic)	-0.1496	(0.0762)
birth in the curr. period	-0.3346	(0.0966)
female wage		
wage intercept	2.8585	(0.0153)
returns to schooling (linear)	-0.2003	(0.0116)
returns to schooling (quadratic)	0.1944	(0.0057)
returns to experience (linear)	-0.0372	(0.0097)
returns to experience (quad)	0.0244	(0.0018)
male earnings		
wage intercept	2.7403	(0.0174)
returns to schooling (linear)	-0.0303	(0.0011)
returns to schooling (quadratic)	0.0029	(0.0000)
returns to experience (linear)	0.0495	(0.0014)
returns to experience (quad)	-0.0007	(0.0000)

parameter estimates - cont'd	Value	St.dev.
error structure		
$\sigma_{leisure}$	0.9581	(0.0113)
$\sigma_{fertility}$	2.8378	(3.4351)
$\sigma_{childcare}$	2.2825	(0.9763)
$\sigma_{wage.wife}$	0.2487	(0.0068)
$\sigma_{wage.wife.meas.error}$	0.1841	(0.0075)
$\sigma_{wage.husband}$	0.218	(0.0011)
number of observations		5000
LL :		6104.2246

Table 5.5: Comparison of observed and predicted moments

	observed moments	baseline prediction
employment rate	0.835	0.829
fertility rate per period	0.067	0.068
child care use rate	0.326	0.384
women's earnings - mean	21.767	22.936
- st. dev.	6.311	7.188
men's earnings - mean	36.92	36.231
- st. dev.	9.724	11.698

Figure 5.10: Comparison of observed and predicted dynamics within the main variables of interest over the period 2002-2009, plots of initial levels and subsequent changes

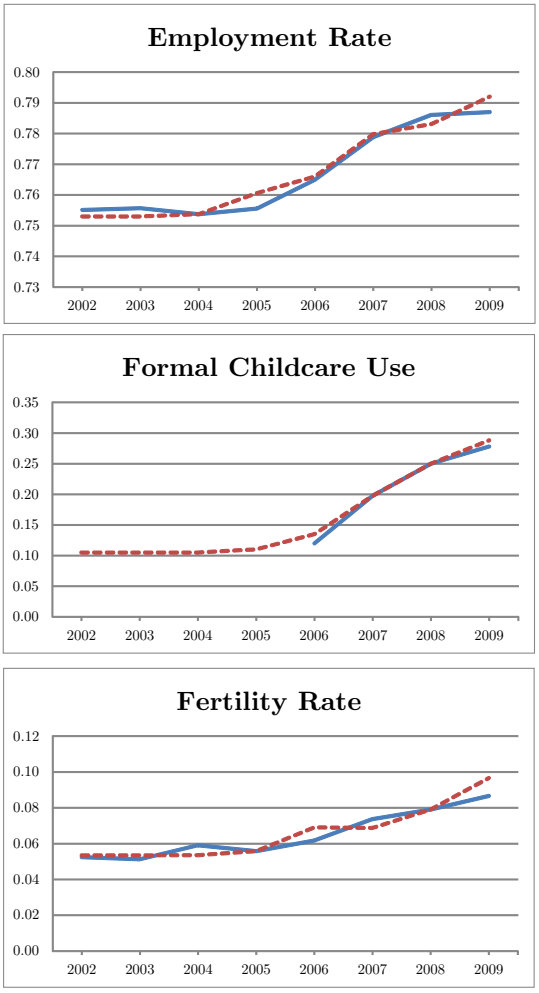
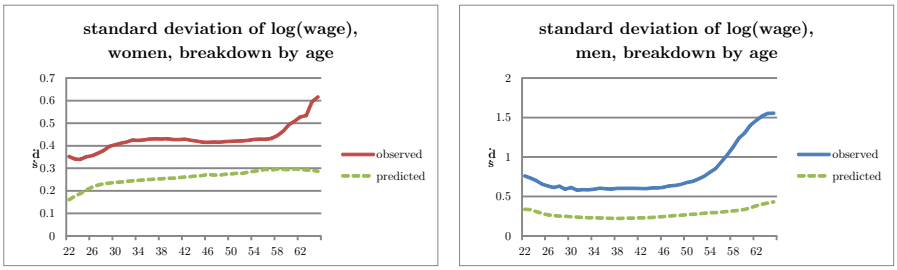


Figure 5.11: Comparison of observed and predicted earnings spreads for women and men living with a partner, breakdown by age of the worker



early retirement decisions which could influence the composition of the older workforce and lead to higher wage dispersion.

5.5.3 *Counterfactual simulations*

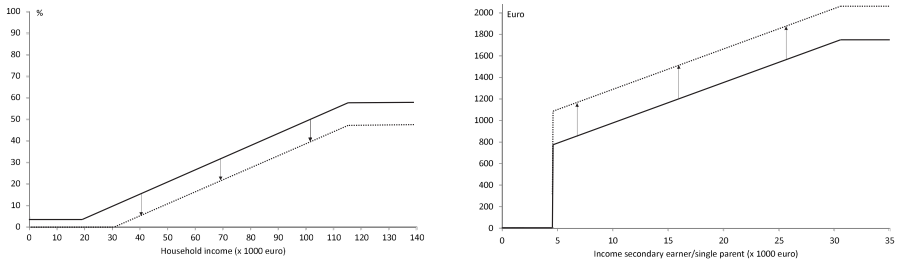
I use the obtained estimates of structural parameters to study the effectiveness of different types of fiscal stimuli for working mothers. I focus on two policies in particular, assessing the relative effectiveness of the implemented reforms of child care subsidies and in-work tax credits in the Netherlands. I derive the reform's effects on labour participation, hours worked and government finances, so that I could quantify how much money has to be spent under each of the reform scenarios in order to induce a comparable increase of female labor supply. The analysis is conducted by performing counterfactual simulations, letting an artificial sample of women be subject to changes in the one of the two policy regimes. I quantify women's response to the policy change relative to the baseline scenario, and calculate the corresponding effect on government budget.

The baseline scenario mimics the tax & subsidy system which was in place in 2009.[24] The child care reform scenario takes the baseline child care subsidy system and lowers the parental contribution rate by 12 percentage points. The IWTC reform scenario increases the 2009 levels of credit by 318 Euro for all female workers with earnings above the minimal income threshold. The baseline and reformed policy rates for both scenarios are presented in the Figure 5.12. The reformed policy rates are chosen to satisfy ex-ante cost equivalence - under the assumption that child care use and labor participation remains at the 2009 levels, both reforms would lead to an increase in public spending of 100 million Euro. This cost equivalence is bound to be broken by behavioral shifts induced by the policy changes, but it serves as a good

24 I have also performed policy simulations with alternative baseline tax systems, but the results remained stable.

reference point for counterfactual simulations of reforms with costs of similar magnitude.

Figure 5.12: The effects of simulated reforms on the corresponding policy rates



I maintain that the policy changes are unexpected by the households, and that they assume the new policy regimes to remain in place indefinitely.[25] Table 5.6 shows the impact of the two policy reforms on women's choices. The results documented in the table represent the 'short-run' effects, corresponding to the impact in period immediately following the policy change. They are expressed as percentage point changes in the variables of interest.

Table 5.6: Effectiveness of fiscal stimuli of 100 million Euro, 1 year from the introduction

	Child care subsidy	EITC
Intended costs	100.0	100.0
Realized SR costs	**155.6**	**92.6**
Labor supply effect	+1.29%	+0.72%
Child care demand	+13.95%	+1.45%
Probability of childbirth	+2.15%	+1.08%
Cost of raising LS by 0.1%	12.1 mil.	12.8 mil.
Cost of raising PoC by 0.1%	7.2 mil.	8.6 mil.

In all accounts, the reforms are shown to have positive effects on the choice variables. The female labor supply is enhanced more by the child care subsidy reform, but this advantage is counterbalanced

25 The households' expectations that the reforms remain in place indefinitely can be replaced by alternative forms of expectations, *e.g.*, that the reforms are only temporary and that the family tax policy will revert to its initial setup after several periods. In such a setting, the reforms are likely to induce lesser labor supply responses, because the households know that mother's investment into human capital will bring less income in the future (compared to the persistent reform setting).

by steep increase of demand for formal child care. The additional child care demand drives up the realized costs of child care reform, because new child care users were not accounted for in the initial cost projection. The spike of child care costs is not observed under the IWTC reform scenario, where child care demand stays relatively low. This result suggests that the forgone maternal time spent with children is substituted to a large extent by informal child care sources. Since the formal child care exhibits lesser response to the changes in IWTC, the realized costs of the reform stay near the initial target. In fact, the costs prove to be lower than expected due to the additional revenue from income tax receipts filed by the women who decided to enter the post-reform labor market.

Last row of Table 5.6 shows another interesting feature of the results. That is, if I quantify the costs of raising labour supply by a fixed proportion (here, ten basis points), the two reforms will prove to require almost exactly the same level of spending. It can be therefore claimed that, in the short-run, the two implemented reforms prove similarly cost-effective.

This equivalence is however bound to be broken in the long-run, since the reforms proved to have dissimilar effects on the other variables of interest. Accordingly, we will observe gradual change of the behavioral effects, which is driven by the dynamic features of the structural model.

The long-run effects of the reform are likely to be different from the short-run effects, because the decisions taken five or ten years from the implementation of the reform will reflect not only the new policy regime, but also the decisions which were influenced by the reform and preceded the period in question. One of the most important considerations in this respect is that an immediate increase in the woman's labor supply is going to translate into additional experience in the later stages of her life. Increased experience will make the woman's labor participation more appealing in years to come because the experience will raise her potential labor earnings. Therefore, we can expect the labor supply effects to increase further in the years after the implementation due to the synergic effect of the new policy regime and the accumulation of women's human capital. This effect is documented in Table 5.7, which shows the simulated behavioral responses quantified 10 years after the implementation of the reforms. We see that the labor supply effects are higher than the ones which are observed immediately. Furthermore, we see that the other effects get more pronounced as well. The share of childless families in the population falls dramatically, and people are predicted to use more formal child care, which is driven by increased labor supply (and higher disposable incomes).

The increasing fertility proves to have important effects on the effectiveness of child care reform. Since the families are facing higher

Table 5.7: Effectiveness of fiscal stimuli of 100 million Euro, 10 years from the introduction

	Child care subsidy	EITC
Labor supply	+2.43%	+1.56%
Child care demand	+20.15%	+3.11%
Share of families with children	+11.4%	+6.39%
Yearly costs	**217.4**	**82.6**
Cost of raising LS by 0.1%	9.0 mil.	5.3 mil.
Cost of raising PoC by 0.1%	10.1 mil.	7.6 mil.

probability of having a baby in each year, the pool of child care eligible children will steadily expand as the cohorts of post-reform babies grow older (the children remain eligible for child care subsidies until the age of twelve). As a result, stimulated fertility will considerably increase long-run costs of the reform. This outcome is in sharp contrast with the long-run results corresponding to the IWTC reform. After 10 years, the maintenance costs of the reform are predicted to decline further, following the additional increase of tax revenue coming from larger female workforce.

The reform comparison above exposes hidden costs of the child care subsidy reform, which are not to be realized in the initial post-reform years. These costs correspond to reform-induced fertility increase, an impact which is most likely unintended by the policy makers. For that reason, it may well escape their awareness, and therefore also their cost projections. As I have shown, such costs are non-trivial. They are capable of erasing significant portion of tax revenues, and they proved to reverse the policy recommendations based on the short-run effects of the policies. This finding bolsters the presumption that a sound policy analysis should look beyond the traditional metric of key variable of interest, but also scrutinize the second-order effects which are likely to be induced by the policy change, and which have the potential to influence government's budget.

5.6 CONCLUSION

This paper develops a dynamic structural model of decision making of married and cohabiting women in the Netherlands. The unique feature of the model is that it allows the households to evaluate three interdependent decisions: the woman's decision to work, the fertility

decision and (conditional on having children) the decision to use formal child care. Each of the decisions considered has been found to be responsive to the changes in financial incentives and they were shown to play an important role in evaluating cost-effectiveness of fiscal stimuli for working mothers.

The model is estimated using a rich administrative panel covering years 2001-2009, which provides high-quality information about individual labour market histories and child care expenditures. Identification is aided by exploiting two large-scale reforms of Dutch family policy which raised the government spending on In-Work Tax Credits and Child care subsidies. Resulting variation in the tax and subsidy rates is explicitly incorporated in the budget constraint.

The model fits the data well, and is capable of replicating the observed dynamics in the variables of interest. The optimized structural parameters are used for comparative analysis of the child care subsidy reform and the IWTC reform. The results show that the two policy reforms induce similarly strong changes in terms of maternal labor supply, as well as short-run government revenue. The fertility is stimulated more by the child care subsidy reform. However, the equivalence of budgetary costs is broken in the long run. Reform-induced increase in fertility rates raises the costs of the child care reform above its short-run projection, as more children become eligible for subsidized child care in the years after the implementation. This finding should be of particular interest for policy makers, since the costs of reform related to the fertility are not to be fully realized sooner than several years after the implementation.

The question which remains to be answered is which of the policies fares better in terms of reaching the two policy goals - to promote female labor participation, and to increase fertility rates. I have shown that the IWTC reform is more cost-efficient in terms of labor supply stimulation, however it falls short of the fertility effect of child care subsidy reform. The two policies should be therefore evaluated with these limitations in mind, and the policy makers should decide for one or another, depending on their priorities - be it budgetary temperance, or fertility of the populace.

A number of improvements and extensions can be considered. Firstly, the error structure can be made more complex, either by allowing for contemporaneous dependence across choice-specific error terms, or by modeling temporal dependence of individual errors. Secondly, the choice variables in the utility function could be further interacted with observable characteristics, allowing for more comprehensive treatment of observable heterogeneity. The treatment of unobservable heterogeneity could be made more flexible by allowing for latent types, similar to the ones presented in Chapters 4 & 3. The set of decisions faced by the household could be also extended, allow-

ing for the responses of the male partner and household formation decisions.

5.7 ACKNOWLEDGMENTS

I have benefited from comments and suggestions by Jaap Abbring, Flavio Cunha, Eric French, David Green, Peter Haan, Egbert Jongen, Guyonne Kalb, Janneke Plantenga, David Ribar, Martin Salm, Andrew Shephard, Arthur van Soest, Petra Todd and Ken Wolpin. Remaining errors are my own.

BIBLIOGRAPHY

Aaberge, R., U. Colombino, and S. Strom (1999): "Labour Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints," *Journal of Applied Econometrics*, 14(4), 403–22.

Aaberge, R., J. K. Dagsvik, and S. Strom (1995): " Labor Supply Responses and Welfare Effects of Tax Reforms," *Scandinavian Journal of Economics*, 97(4), 635–59.

Akgunduz, Y. E., and J. Plantenga (2013): "Competition for a better future? Effects of competition on child care quality," Working Papers, Utrecht School of Economics 13-14, Utrecht School of Economics.

Alesina, A., A. Ichino, and L. Karabarbounis (2011): "Gender-Based Taxation and the Division of Family Chores," *American Economic Journal: Economic Policy*, 3(2), 1–40.

Apps, P., J. Kabátek, R. Rees, and A. Van Soest (2012): "Labor supply heterogeneity and demand for child care of mothers with young children," IZA Discussion Paper 7007, Bonn.

Apps, P., and R. Rees (1988): "Taxation and the household," *Journal of Public Economics*, 35(3), 355–369.

——— (1999): "On the taxation of trade within and between households," *Journal of Public Economics*, 73(2), 241–263.

——— (2009): *Public Economics and the Household*. Cambridge University Press, Cambridge.

——— (2011): "Optimal Taxation and Tax Reform for Two-Earner Households," *CESifo Economic Studies*, 57(2), 283–304.

Baker, M., J. Gruber, and K. Milligan (2008): "Universal Child Care, Maternal Labor Supply, and Family Well-Being," *Journal of Political Economy*, 116(4), 709–745.

Bargain, O., K. Orsini, and A. Peichl (2014): "Comparing Labor Supply Elasticities in Europe and the United States: New Results," *Journal of Human Resources*, 49(3), 723–838.

Baughman, R., and S. Dickert-Conlin (2003): "Did Expanding the EITC Promote Motherhood?," *American Economic Review*, 93(2), 247–251.

Berlinski, S., and S. Galiani (2007): "The effect of a large expansion of pre-primary facilities on preschool attendance and maternal employment," *Labour Economics*, 14, 665–680.

Bernal, R. (2008): "The Effect of Maternal Employment and Child Care on Children's Cognitive Development," *International Economic Review*, 9(4), 1173–1209.

Bernal, R., and M. P. Keane (2011): "Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers," *Journal of Labor Economics*, 29(3), 459 – 512.

Bettendorf, L., E. Jongen, and P. Muller (2012): "Childcare subsidies and labour supply: evidence from a large Dutch reform," CPB Discussion Paper 217, The Hague.

Blau, D. (2003): "Child care subsidy programs," in *Means-Tested Transfer Programs in the United States*, ed. by R. Moffitt, pp. 443–516. NBER.

Blau, D., and J. Currie (2006): *Handbook of the Economics of Education*chap. Preschool, day care, and after school care: who's minding the kids?, pp. 1163–1278. Elsevier.

Blau, D. M., and A. P. Hagy (1998): "The Demand for Quality in Child Care," *Journal of Political Economy*, 106(1), 104–146.

Blau, F. D. (2012): *Gender, Inequality, and Wages*, no. 9780199665853 in OUP Catalogue, Oxford University Press. Oxford University Press.

Bloemen, H., and E. Stancanelli (2014): "Market hours, household work, child care, and wage rates of partners: an empirical analysis," *Review of Economics of the Household*, 12(1), 51–81.

Blundell, R. (2012): "Tax policy reform: the role of empirical evidence," *Journal of the European Economic Association*, 10(1), 43–77.

Blundell, R., A. Bozio, and G. Laroque (2013): "Extensive and Intensive Margins of Labour Supply: Work and Working Hours in the US, the UK and France," *Fiscal Studies*, 34(1), 1–29.

Blundell, R., M. C. Dias, C. Meghir, and J. Shaw (2013): "Female Labor Supply, Human Capital and Welfare Reform," Cowles Foundation Discussion Papers 1892R, Cowles Foundation for Research in Economics, Yale University.

Blundell, R., A. Duncan, J. McCrae, and C. Meghir (2000): "The labour market impact of the Working Families' Tax Credit," *Fiscal Studies*, 21(1), 75–104.

Blundell, R., A. Gosling, H. Ichimura, and C. Meghir (2007): "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds," *Econometrica*, 75(2), 323–363.

Blundell, R., T. MaCurdy, and C. Meghir (2007): "Labor Supply Models: Unobserved Heterogeneity, Nonparticipation and Dynamics," in *Handbook of Econometrics*, ed. by J. Heckman, and E. Leamer, vol. 6 of *Handbook of Econometrics*, chap. 69. Elsevier.

Blundell, R., and A. Shephard (2012): "Employment, hours of work and the optimal taxation of low income families," *Review of Economic Studies*, 79(2), 481–510.

Blundell, R., and I. Walker (1986): "A Life-Cycle Consistent Empirical Model of Family Labour Supply Using Cross-Section Data," *Review of Economic Studies*, 53(4), 539–58.

Bourguignon, F., and T. Magnac (1990): "Labor Supply and Taxation in France," *Journal of Human Resources*, 25(3), 358–389.

Brewer, M., and J. Browne (2006): "The effect of the Working Families' Tax Credit on labour market participation," IFS Briefing Note no. 69.

Brewer, M., M. Francesconi, P. Gregg, and J. Grogger (2009): "Feature: in-work benefit reform in a cross-national perspective - introduction," *Economic Journal*, 119, F1–F14.

Brewer, M., A. Ratcliffe, and S. dSmith (2012): "Does welfare reform affect fertility? Evidence from the UK," *Journal of Population Economics*, 25(1), 245–266.

Burda, M., D. Hamermesh, and P. Weil (2013): "Total work and gender: facts and possible explanations," *Journal of Population Economics*, 26(1), 239–261.

Burda, M., M. Harding, and J. Hausman (2008): "A Bayesian mixed logit & probit model for multinomial choice," *Journal of Econometrics*, 147(2), 232–246.

Callan, T., A. van Soest, and J. R. Walsh (2009): "Tax Structure and Female Labour Supply: Evidence from Ireland," *LABOUR*, 23(1), 1–35.

Cascio, E. (2009): "Maternal labor supply and the introduction of kindergartens into American public schools," *Journal of Human Resources*, 44(1), 140–170.

Chetty, R., A. Looney, and K. Kroft (2009): "Salience and taxation: theory and evidence," *American Economic Review*, 99(4), 1145–1177.

Claeskens, G., and N. L. Hjort (2003): "The Focused Information Criterion," *Journal of the American Statistical Association*, 98(464), pp. 900–945.

CONNELLY, R. (1992): "The Effect of Child Care Costs on Married Women's Labor Force Participation," *The Review of Economics and Statistics*, 74(1), 83–90.

CONNELLY, R., AND J. KIMMEL (2003): "Marital status and full-time/part-time work status in child care choices," *Applied Economics*, 35(7), 761–777.

DAGSVIK, J. K. (1994): "Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes," *Econometrica*, 62(5), 1179–1205.

DE BOER, H.-W., E. JONGEN, AND J. KABÁTEK (2014): "The effectiveness of fiscal stimuli for working parents," CPB Working Paper 116.

DEATON, A., AND C. PAXSON (1994): "Intertemporal Choice and Inequality," *Journal of Political Economy*, 102(3), 437–467.

DEL BONO, E., A. WEBER, AND R. WINTER-EBMER (2012): "Clash of career and family: Fertility decisions after job displacement," *Journal of the European Economic Association*, 10(4), 659–683.

DOIRON, D., AND G. KALB (2005): "Demands for Child Care and Household Labour Supply in Australia," *The Economic Record*, 81(254), 215–236.

DONALD, S. G., AND D. S. HAMERMESH (2009): "A structural model of the fixed time costs of market work," *Economics Letters*, 104(3), 125–128.

EISSA, N., AND H. W. HOYNES (2004): "Taxes and the labor market participation of married couples: the earned income tax credit," *Journal of Public Economics*, 88(9-10), 1931–1958.

EUROPEAN UNION (2014): *Taxation trends in the European Union, 2014 edition*. European Union.

EUWALS, R., M. KNOEF, AND D. VAN VUUREN (2011): "The trend in female labour force participation: what can be expected for the future?," *Empirical Economics*, 40, 729–753.

FITZPATRICK, M. (2010): "Preschoolers enrolled and mothers at work? The effects of universal prekindergarten," *Journal of Labor Economics*, 28(1), 51–85.

——— (2012): "Revising our thinking about the relationship between maternal labor supply and preschool," *Journal of Human Resources*, 47(3), 583–612.

FRANCESCONI, M. (2002): "A Joint Dynamic Model of Fertility and Work of Married Women," *Journal of Labor Economics*, 20(2), 336–380.

FRAZIS, H., AND J. STEWART (2012): "How to Think About Time-Use Data : What Inferences Can We Make About Long and Short-Run Time Use from Time Diaries ?," *Annales d'Economie et de Statistique*, (105-106), 11.

GELBER, A. M., AND J. W. MITCHELL (2012): "Taxes and Time Allocation: Evidence from Single Women and Men," *Review of Economic Studies*, 79(3), 863–897.

GEYER, J., P. HAAN, AND K. WROHLICH (2014): "The effects of family policy on mothers' labor supply: combining evidence from a structural model and a natural experiment," SOEPpapers on Multidisciplinary Panel Data Research, Berlin.

GONG, X., AND R. BREUNIG (2011): "Estimating Net Child Care Price Elasticities of Partnered Women With Pre-School Children Using a Discrete Structural Labour Supply-Child Care Model," CEPR Discussion Papers 653, Centre for Economic Policy Research, Research School of Economics, Australian National University.

GONG, X., AND R. BREUNIG (2012): "Child care assistance: are subsidies or tax credits better?," IZA Discussion Paper 6606, Bonn.

GONG, X., AND A. VAN SOEST (2002): "Family structure and female labor supply in Mexico City," *Journal of Human Resources*, 37(1), 163–191.

HAAN, P., AND K. WROHLICH (2011): "Can child care policy encourage employment and fertility?: Evidence from a structural model," *Labour Economics, Elsevier*, 18(4), 498–512.

HANSEN, J., AND X. LIU (2011): "Estimating labor supply responses and welfare participation: using a natural experiment to validate a structural model," IZA Discussion Paper 5718, Bonn.

HAVNES, T., AND M. MOGSTAD (2011): "Money for nothing? Universal child care and maternal employment," *Journal of Public Economics*, 95, 1455–1465.

HAVNES, T., AND M. MOGSTAD (2014): "Is universal child care leveling the playing field?," *Journal of Public Economics*, (0), –.

HECKMAN, J., AND B. SINGER (1984): "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica*, 52, 271–320.

HOTZ, V., C. MULLIN, AND J. SCHOLZ (2010): "Examining the effect of the Earned Income Tax Credit on the labor market participation of families on welfare," mimeo, Duke University.

Hotz, V., and J. Scholz (2003): "The Earned Income Tax Credit," in *Means-Tested Transfer Programs in the U.S.* , ed. by R. Moffitt. University of Chicago Press.

Hoynes, H. W. (1996): "Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation under AFDC-UP," *Econometrica*, 64(2), 295–332.

Immervoll, H., and M. Pearson (2009): "A good time for making work pay? Taking stock of in-work benefits and related measures across the OECD," IZA Policy Paper No. 3, Bonn.

Kabátek, J. (2013): "Iteration capping for discrete choice models using the EM algorithm," Discussion Paper 2013-019, Tilburg University, Center for Economic Research.

Kabatek, J., A. van Soest, and E. Stancanelli (2014): "Income taxation, labour supply and housework: A discrete choice model for French couples," *Labour Economics*, 27(0), 30 – 43.

Kalenkoski, C. M., D. C. Ribar, and L. S. Stratton (2005): "Parental Child Care in Single-Parent, Cohabiting, and Married-Couple Families: Time-Diary Evidence from the United Kingdom," *American Economic Review*, 95(2), 194–198.

Keane, M., and R. Moffitt (1998): "A Structural Model of Multiple Welfare Program Participation and Labor Supply," *International Economic Review*, 39(3), pp. 553–589.

Keane, M. P., and K. I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76(4), 648–72.

Kleven, H. (2014): "How can Scandinavians tax so much?," *Journal of Economic Perspectives*, forthcoming.

Kornstad, T., and T. Thoresen (2006): "Effects of family policy reforms in Norway: results from a joint labour supply and childcare choice microsimulation analysis," *Fiscal Studies*, 27(3), 339–371.

——— (2007): "A discrete choice model for labor supply and childcare," *Journal of Population Economics*, 20, 781–803.

Lefebvre, P., and P. Merrigan (2008): "Child-care policy and the labor supply of mothers with young children: a natural experiment from Canada," *Journal of Labor Economics*, 26(3), 519–48.

Leuthold, J. H. (1983): "Home production and the tax system," *Journal of Economic Psychology*, 3(2), 145–157.

LOKSHIN, M. (2004): "Household childcare choices and women's work behavior in Russia," *Journal of Human Resources*, 39(4), 1094–1115.

LUNDIN, D., E. MÖRK, AND B. ÖCKERT (2008): "How far can reduced childcare prices push female labour supply?," *Labour Economics*, 15, 647–659.

MACURDY, T., P. GREEN, AND H. PAARSCH (1990): "Assessing empirical approaches for analyzing taxes and labor supply," *Journal of Human Resources*, 25, 415–490.

MCFADDEN, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 57(5), 995–1026.

MEGHIR, C., AND D. PHILLIPS (2010): "Labour supply and taxes," in *Dimensions of Tax Design: The Mirrlees Review*, ed. by J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba. Oxford University Press.

MILLIGAN, K. (2005): "Subsidizing the Stork: New Evidence on Tax Incentives and Fertility," *The Review of Economics and Statistics*, 87(3), 539–555.

MINISTRY OF FINANCE (2010): "Het kind van de regeling, rapport Brede Heroverwegingen 5," Ministry of Finance, The Hague.

MINISTRY OF SOCIAL AFFAIRS AND EMPLOYMENT (2011): "Kinderopvangtoeslag vanaf 2012," Ministry of Social Affairs and Employment, The Hague.

——— (2014): "Cijfers kinderopvang 2013," Ministry of Social Affairs and Employment, The Hague.

MROZ, T. A. (1999): "Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome," *Journal of Econometrics*, 92(2), 233–274.

MUNDLAK, Y. (1978): "On the pooling of time series and cross section data," *Econometrica*, 46(1), 69–85.

OECD (2007): *Babies and Bosses, Reconciling Work and Family Life: A Synthesis of Findings for OECD Countries*. OECD, Paris.

——— (2013): "Labour Force Statistics," .

OECD (2014): "OECD Family Database," Paris.

PACIFICO, D. (2012): "On the role of unobserved preference heterogeneity in discrete choice models of labour supply," .

PLANTENGA, J., Y. WEVER, B. RIJKERS, AND P. DE HAAN (2005): "Arbeidsmarktparticipatie en de kosten van kinderopvang," *Economisch Statistische Berichten*, 4455, 115.

RIBAR, D. C. (1995): "A Structural Model of Child Care and the Labor Supply of Married Women," *Journal of Labor Economics*, 13(3), 558–97.

ROGERSON, R. (2009): "Market Work, Home Work, and Taxes: A Cross-Country Analysis," *Review of International Economics*, 17(3), 588–601.

ROMIJN, G., J. GOES, P. DEKKER, M. GIELEN, AND F. VAN ES (2008): "MIMOSI: Microsimulatiemodel voor belastingen, sociale zekerheid, loonkosten en koopkracht," CPB Document 161, The Hague.

RUUD, P. (1991): "Extensions of estimation methods using the EM algorithm," *Journal of Econometrics*, 49(3), 305–341.

SAEZ, E. (2002): "Optimal income transfer programs: intensive versus extensive labor supply responses," *Quarterly Journal of Economics*, 117(3), 1039–1073.

SORENSEN, E. (1993): "Continuous Female Workers: How Different Are They from Other Women?," *Eastern Economic Journal*, 19(1), 15–32.

STANCANELLI, E. G. F., AND L. S. STRATTON (2014): "Maids, appliances and couples' housework : the demand for inputs to domestic production," .

STATISTICS NETHERLANDS (2012): "Documentatierapport Arbeidsmarktpanel 1999-2009V1," .

STEINER, V., AND K. WROHLICH (2004): "Household Taxation, Income Splitting and Labor Supply Incentives: A Microsimulation Study for Germany," Discussion paper.

TEKIN, E. (2007): "Child Care Subsidies, Wages, and Employment of Single Mothers," *Journal of Human Resources*, 42(2), 453–487.

TODD, P., AND K. WOLPIN (2006): "Assessing the impact of a school subsidy program in Mexico: using a social experiment to validate a dynamic behavioral model of child schooling and fertility," *American Economic Review*, 95(5), 1384–1417.

TRAIN, K. (2003): *Discrete Choice Methods with Simulation*, no. emetr2 in Online economics textbooks. SUNY-Oswego, Department of Economics.

TRAIN, K. (2008): "EM algorithms for nonparametric estimation of mixing distributions," *Journal of Choice Modelling*, 1, 40–69.

TRAIN, K. (2009): *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge.

VAN SOEST, A. (1995): "Structural models of family labor supply: a discrete choice approach," *Journal of Human Resources*, 30(1), 63–88.

VAN SOEST, A., M. DAS, AND X. GONG (2002): "A structural labor supply model with flexible preferences," *Journal of Econometrics*, 107(1-2), 345–374.

VELLA, F., AND M. VERBEEK (1998): "Whose Wages do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men," *Journal of Applied Econometrics*, 13, 163–183.

——— (1999): "Two-step estimation of panel data models with censored endogenous variables and selection bias," *Journal of Econometrics*, 90(2), 239–263.